

NISTIR 8280

**Ongoing Face Recognition
Vendor Test (FRVT)
Part 3: Demographic Effects**

**Annex 17 : Candidate list score magnitudes by sex and
race**

This document is an annex of NIST Interagency Report 8280:
<https://doi.org/10.6028/NIST.IR.8280>

2019/12/19

NIST
**National Institute of
Standards and Technology**
U.S. Department of Commerce

1 Overview

This annex includes figures that show similarity scores returned on candidate lists produced when searching images of asian, black and white men and women against enrolled galleries containing either 1.6 million or 12 million subjects. The figures show the extent to which algorithms produce different scores on different demographic groups.

2 Data

The mugshot dataset and its metadata is described in Annex [Annex 1](#).

The figures in this annex were generated from the four experiments detailed in [Table 1](#).

Note we have insufficient data in these trials to look at effect of age.

3 Plot

Each page contains two heatmaps corresponding to one algorithm. The upper figure shows median non-mate scores for each rank 1 . . . 50. The lower figure shows median mate scores. The medians are computed over the number of searches given in this table.

Kind	Sex	Race	Num. Searches
Nonmated	F	A	1244
Nonmated	F	B	24509
Nonmated	F	W	54183
Nonmated	M	A	3743
Nonmated	M	B	66012
Nonmated	M	W	158845
Mated	F	A	232
Mated	F	B	10926
Mated	F	W	19436
Mated	M	A	1025
Mated	M	B	43322
Mated	M	W	73186

Note the two figures use separate scales - the mate scores are typically higher than the nonmate scores. The figures show median scores, not the tails so they do not convey information about false positives or negatives.

We did not report results for searches where the demographics of the individual were unknown.

Within each figure there are two panels, the top refers to galleries in which subjects are enrolled with a variable number of images - i.e. the lifetime consolidation of all the photographs. The second row refers to galleries in which subjects are enrolled with just one image.

	ENROLLMENT				SEARCH			
	TYPE SEE	POPULATION			MATE		NON-MATE	
	SECTION ??	FILTER	N-SUBJECTS	N-IMAGES	N-SUBJECTS	N-IMAGES	N-SUBJECTS	N-IMAGES
Mugshot trials from enrollment of single images								
1	RECENT	NATURAL	1 600 000	1 600 000				
2	RECENT	NATURAL	12 000 000	12 000 000				
Mugshot trials from enrollment of lifetime images								
3	CONSOL	NATURAL	1 600 000	3 351 206				
4	CONSOL	NATURAL	12 000 000	26 107 917				

Table 1: Enrollment and search sets. Each row summarizes one identification trial. The term “natural” means that subjects were selected without heed to demographics, i.e. in the distribution native to this dataset. The probe images were collected in a different calendar year to the enrollment image. Missing values in rows 2-12 are the same as in row 1.

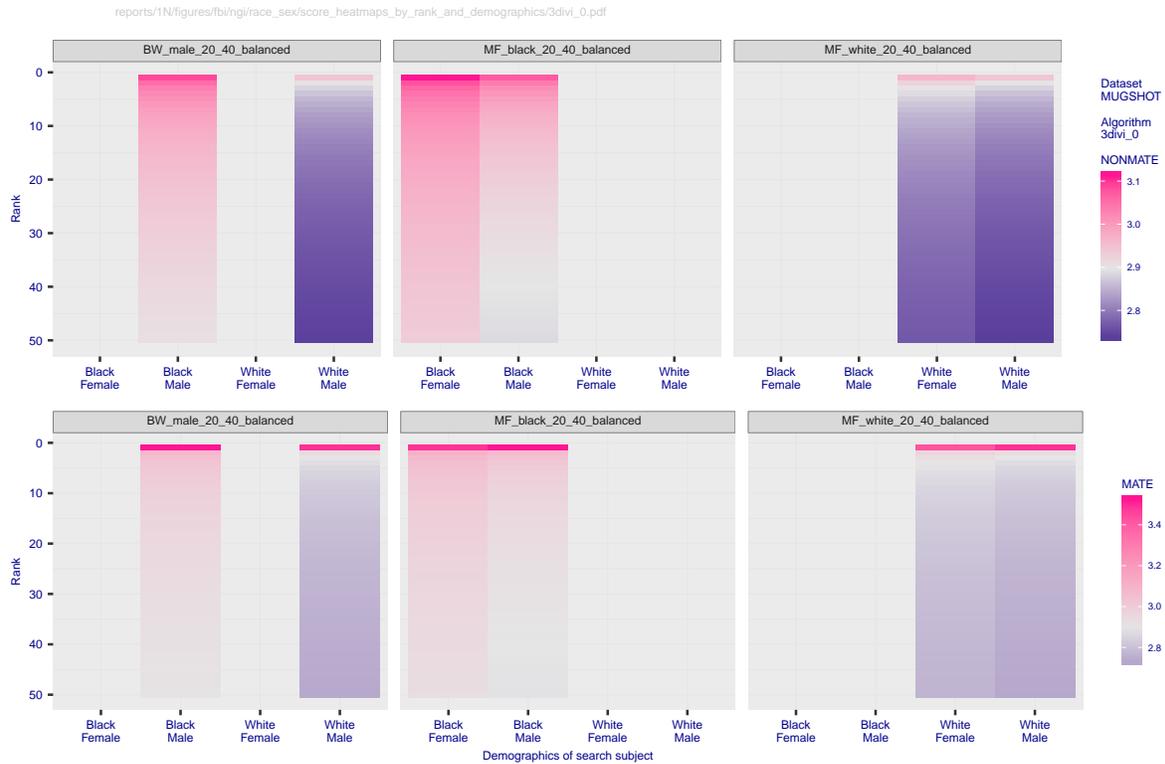


Figure 1: Non-mate score magnitudes by sex and race for mugshot, 3divi-0. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

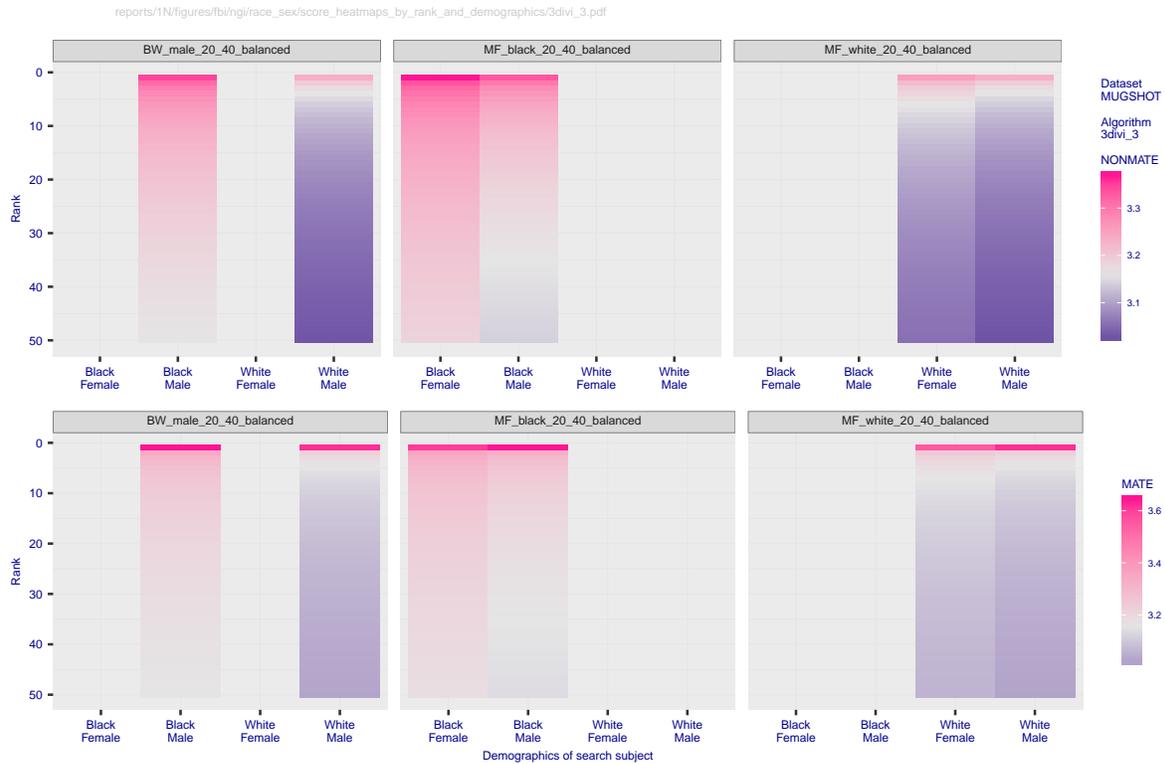


Figure 2: Non-mate score magnitudes by sex and race for mugshot, 3divi-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of N = 1.6 and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

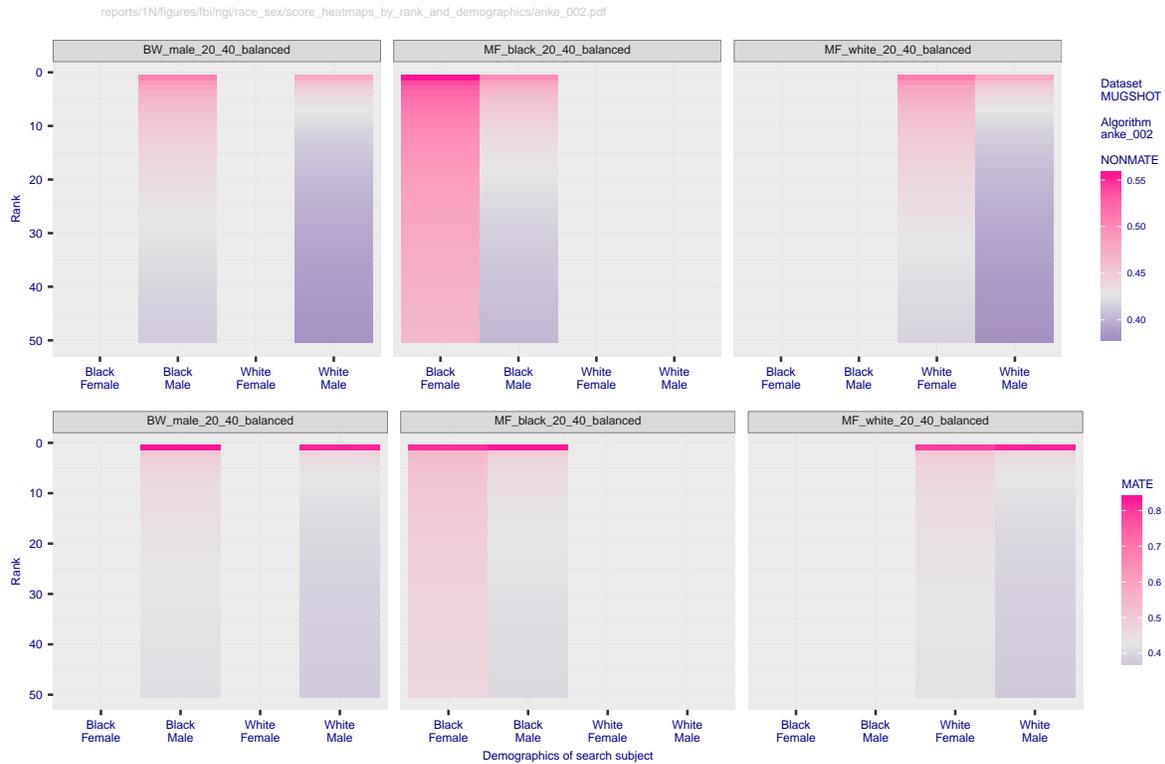


Figure 3: Non-mate score magnitudes by sex and race for mugshot, anke-002. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

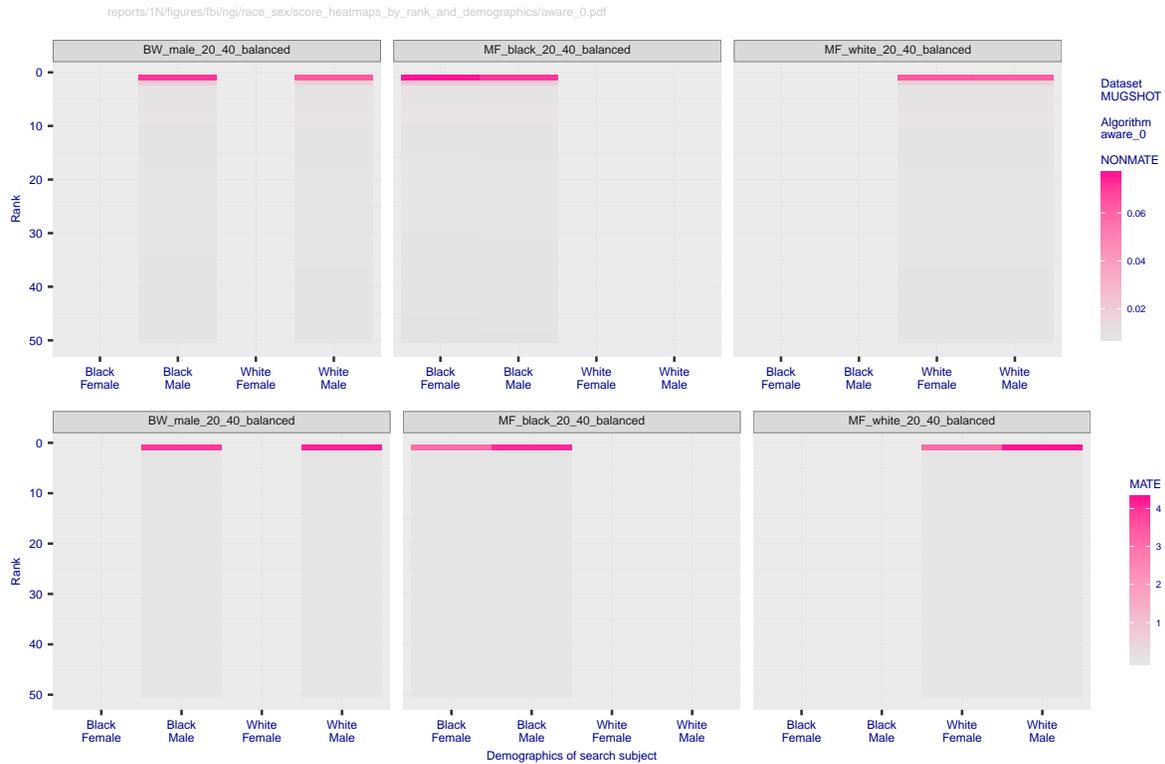


Figure 4: Non-mate score magnitudes by sex and race for mugshot, aware-0. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

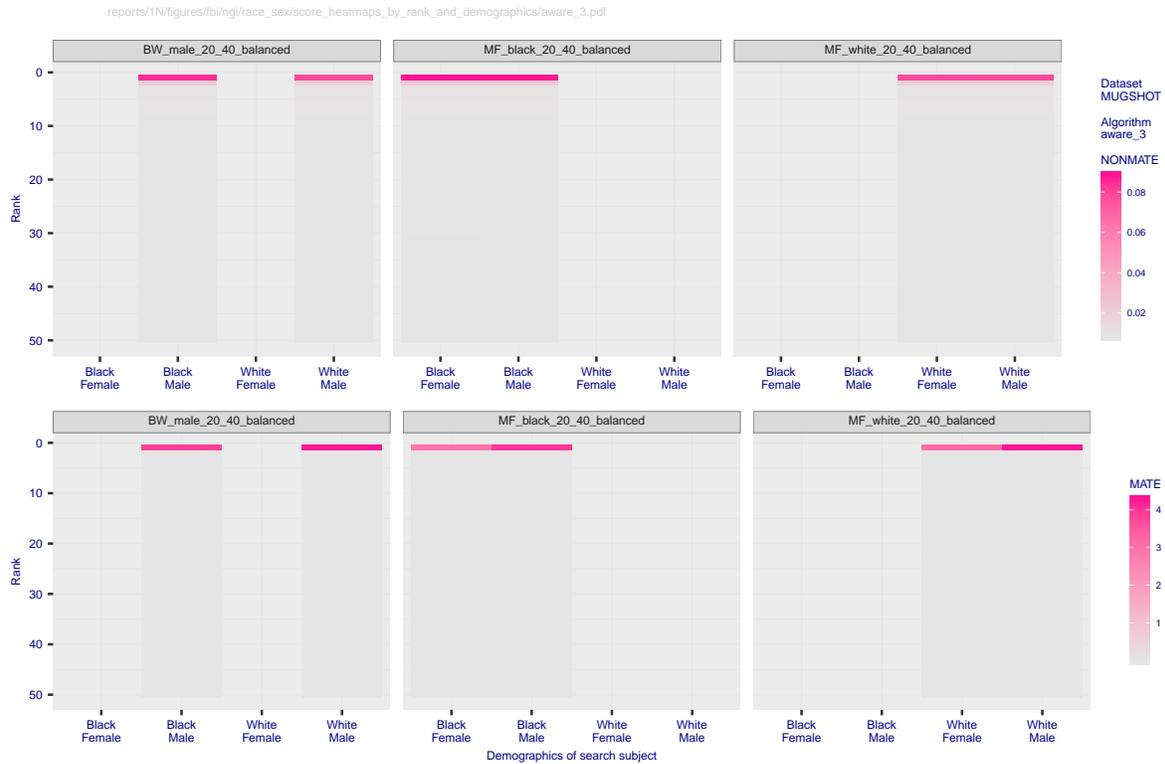


Figure 5: Non-mate score magnitudes by sex and race for mugshot, aware-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

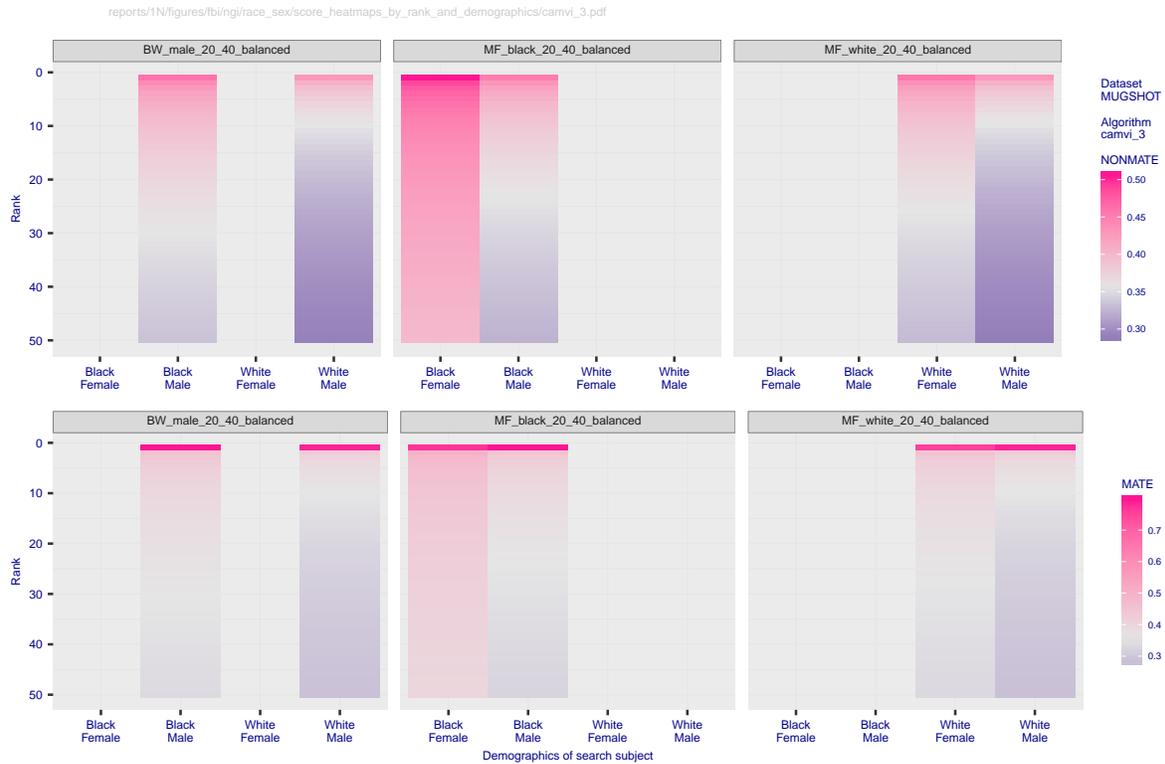


Figure 6: Non-mate score magnitudes by sex and race for mugshot, camvi-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

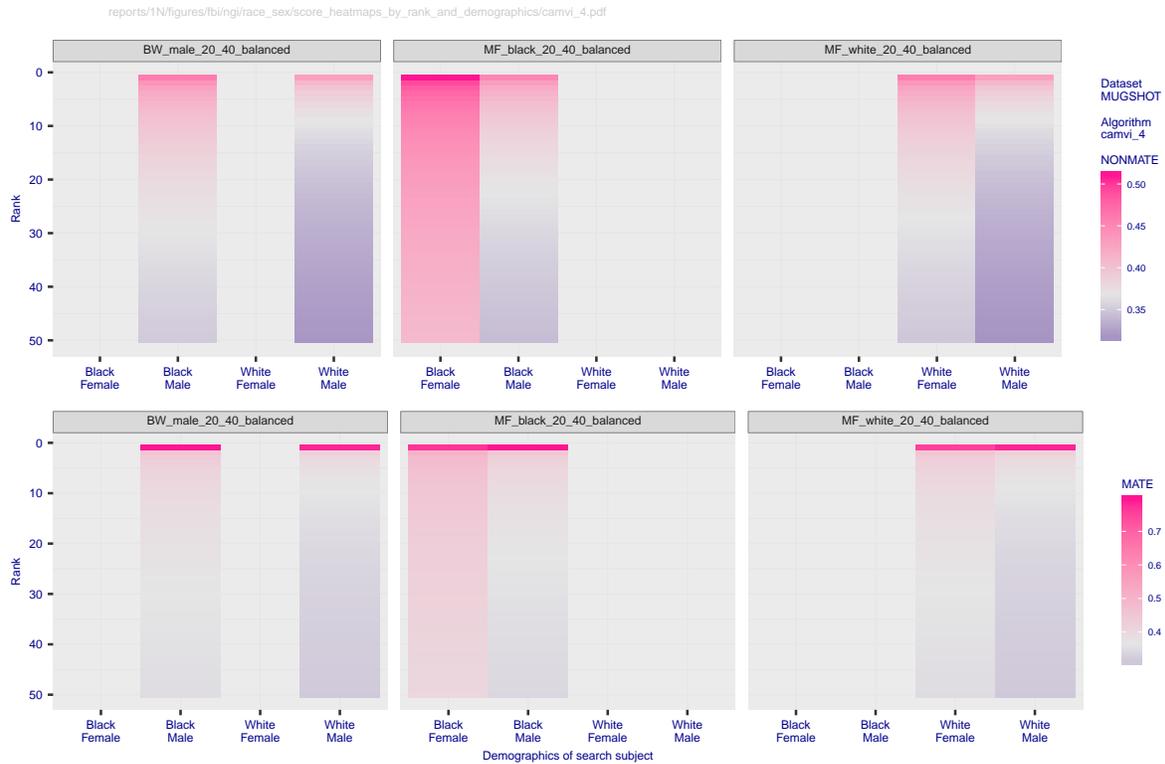


Figure 7: Non-mate score magnitudes by sex and race for mugshot, camvi-4. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

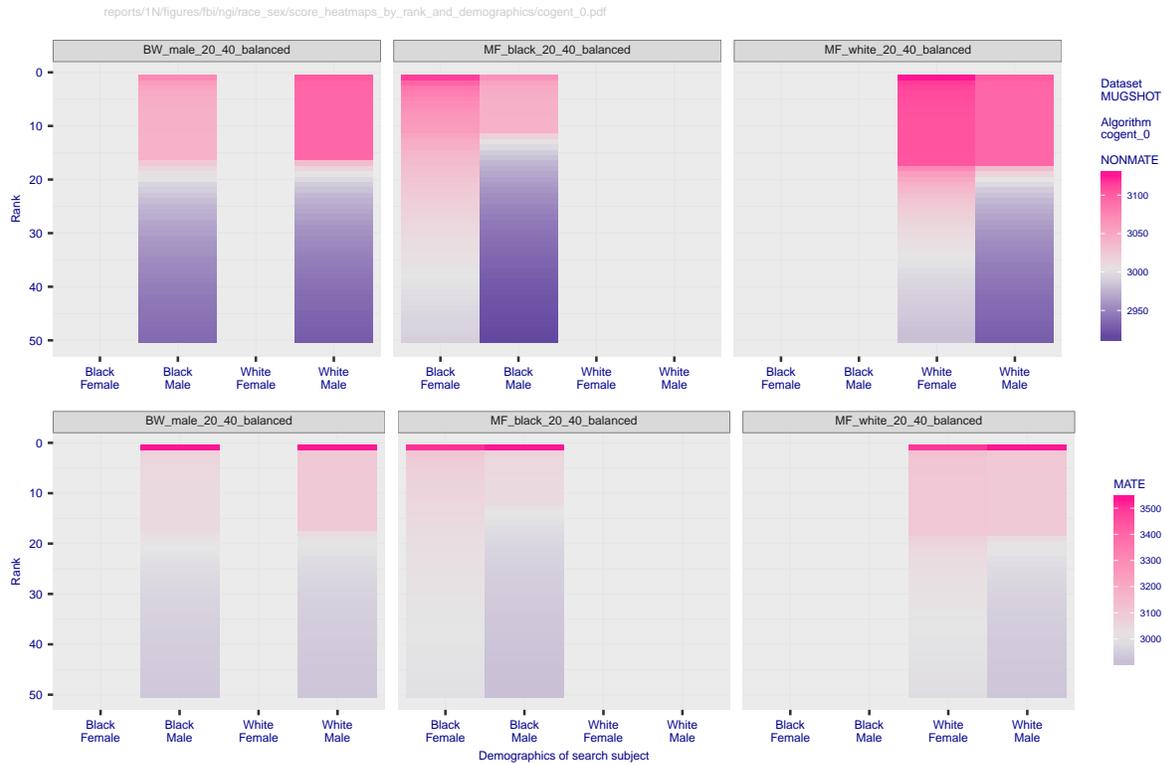


Figure 8: Non-mate score magnitudes by sex and race for mugshot, cogent-0. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

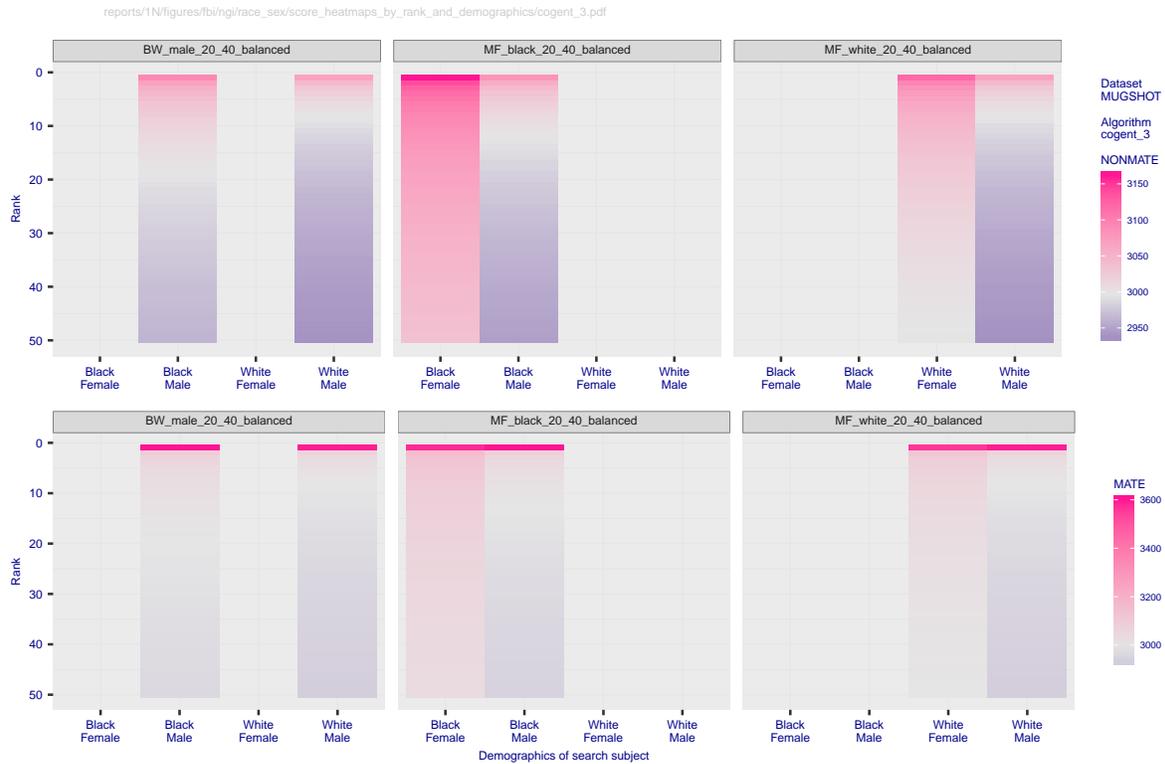


Figure 9: Non-mate score magnitudes by sex and race for mugshot, cogent-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

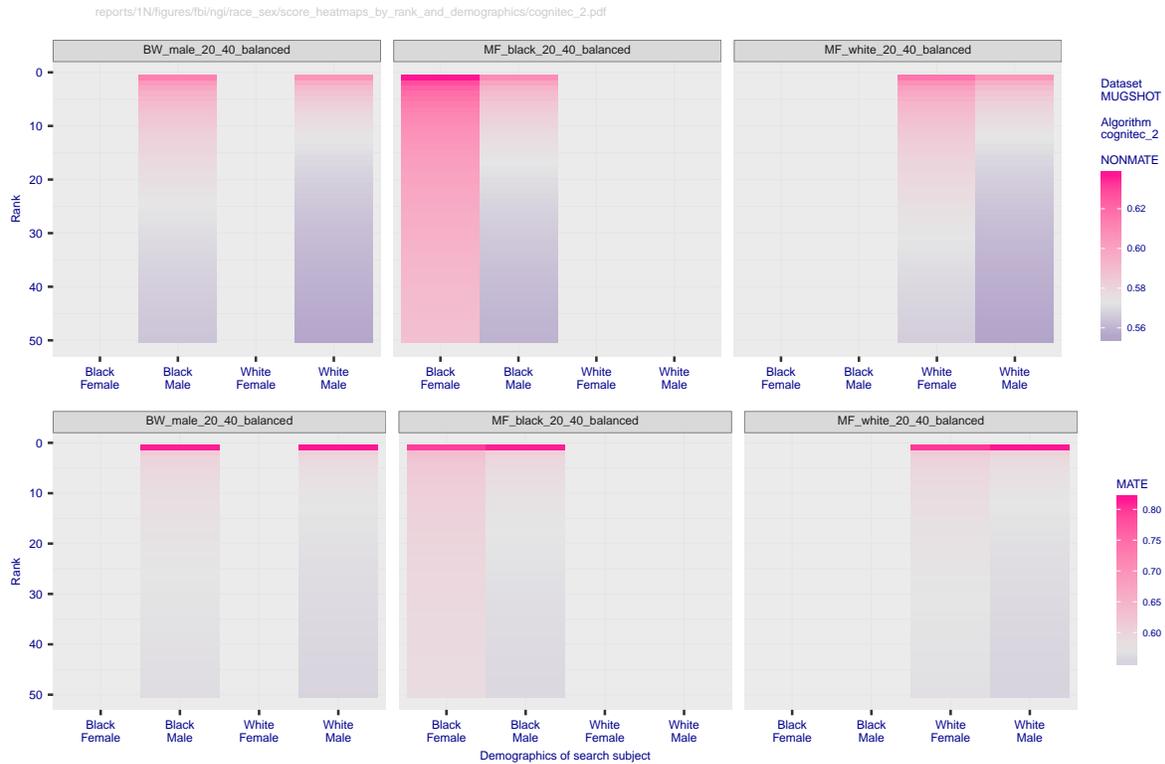


Figure 10: Non-mate score magnitudes by sex and race for mugshot, cognitec-2. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

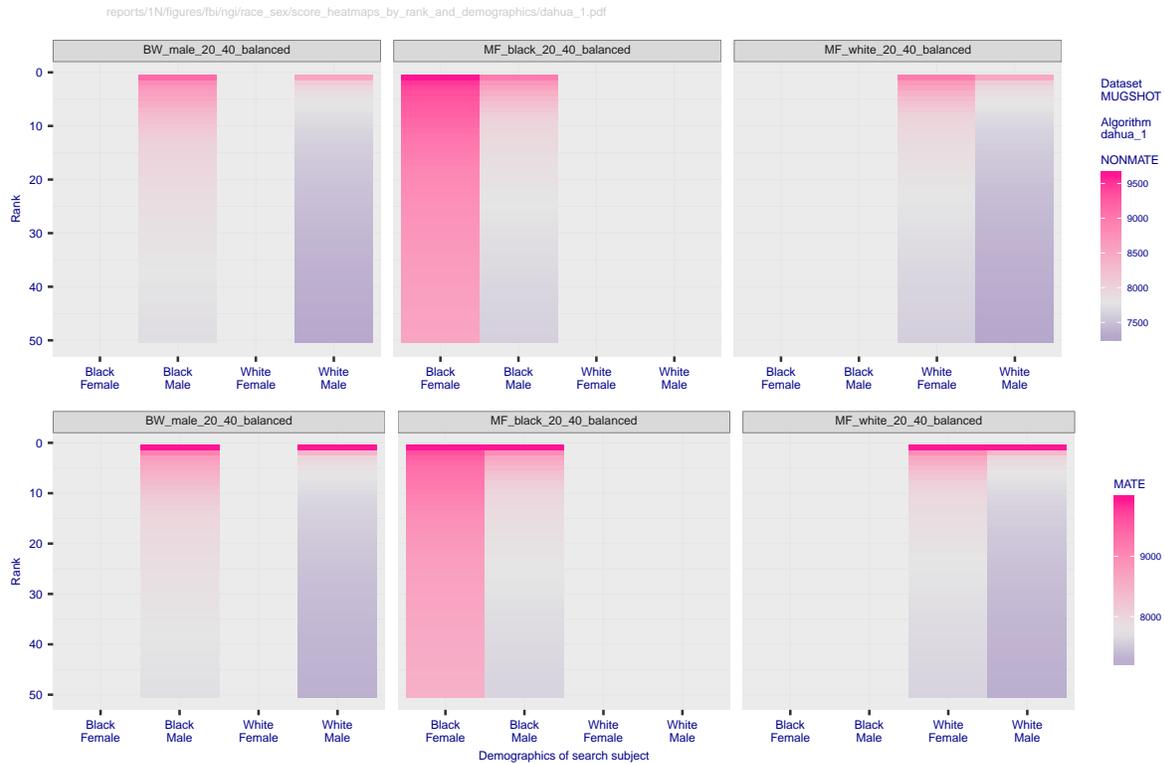


Figure 11: Non-mate score magnitudes by sex and race for mugshot, dahua-1. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

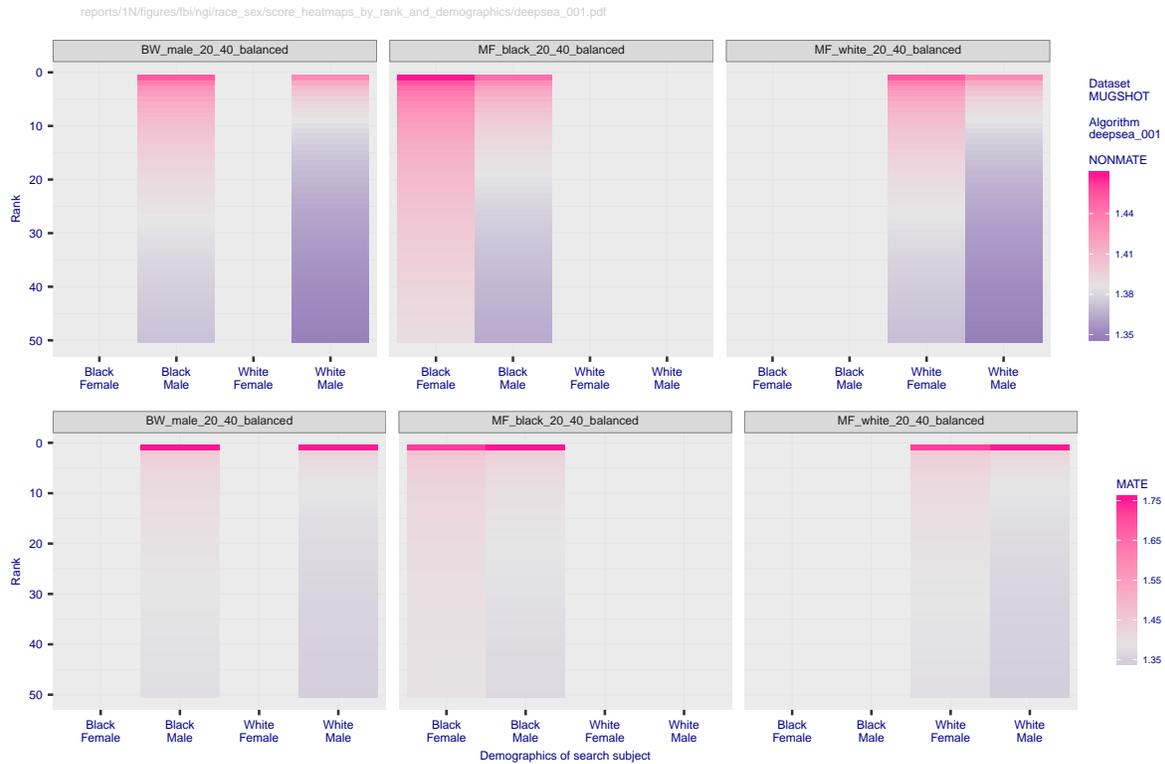


Figure 12: Non-mate score magnitudes by sex and race for mugshot, deepsea-001. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: https://doi.org/10.6028/NIST.IR.8280

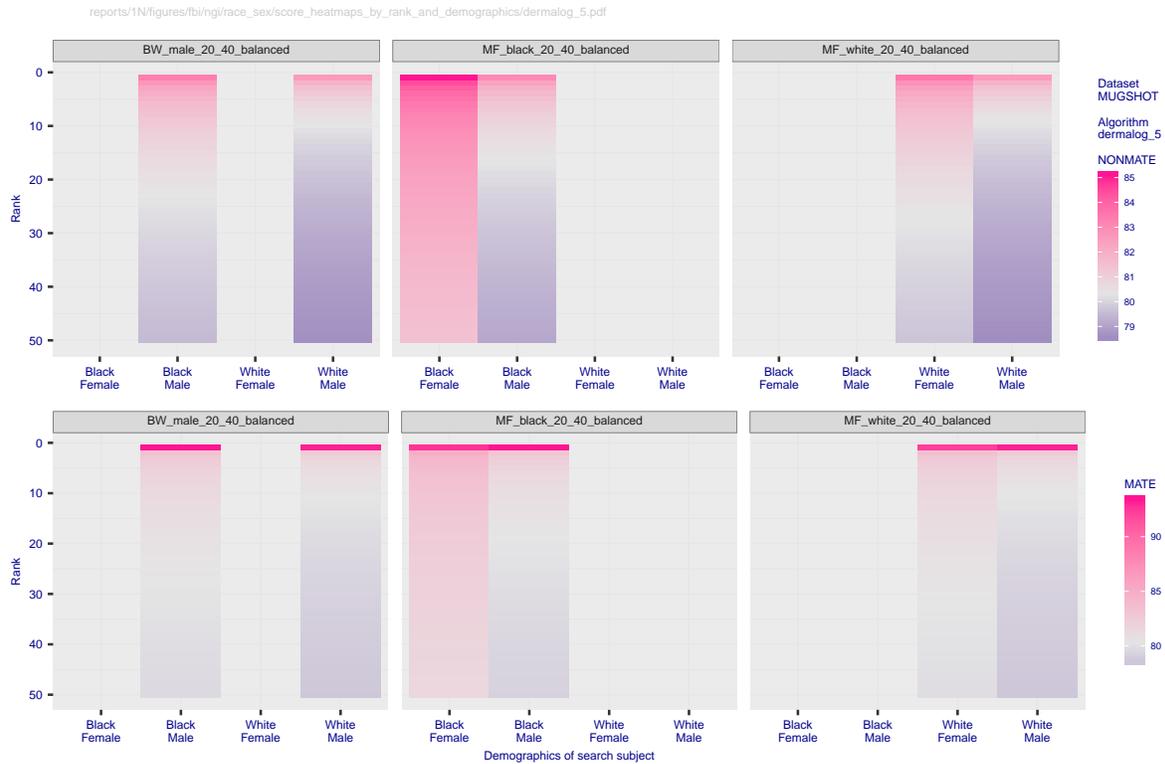


Figure 13: Non-mate score magnitudes by sex and race for mugshot, dermalog-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

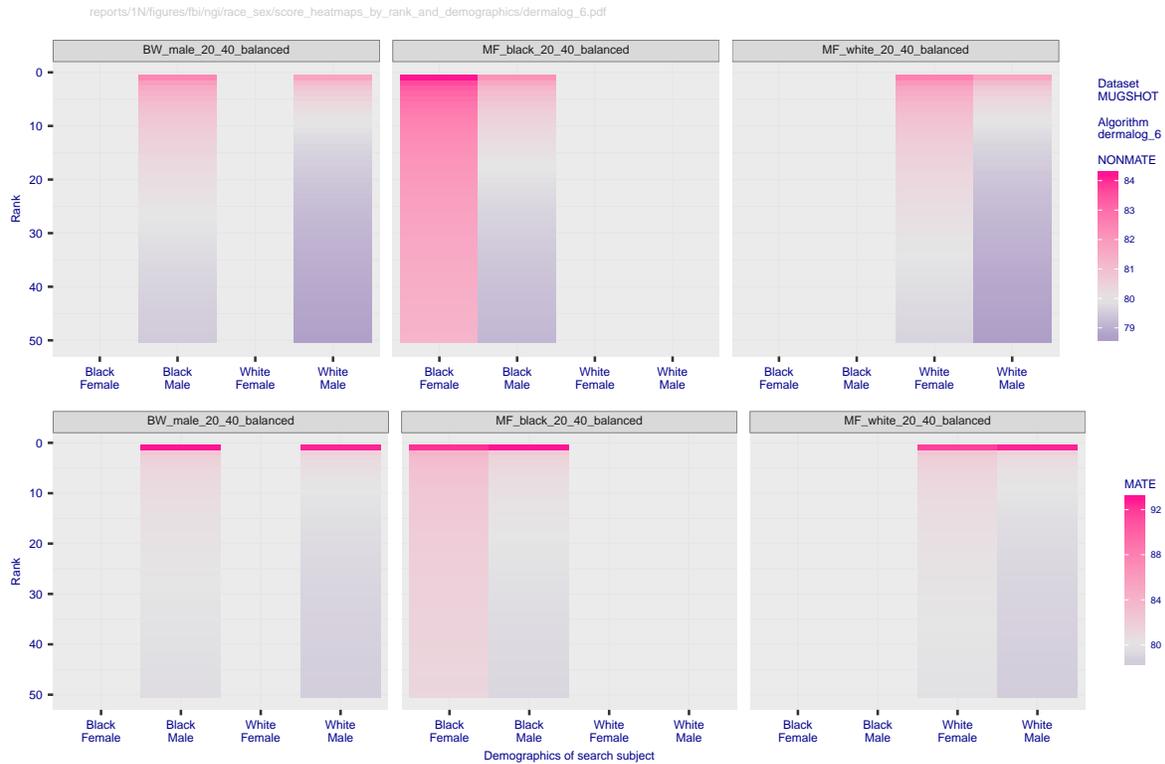


Figure 14: Non-mate score magnitudes by sex and race for mugshot, dermalog-6. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

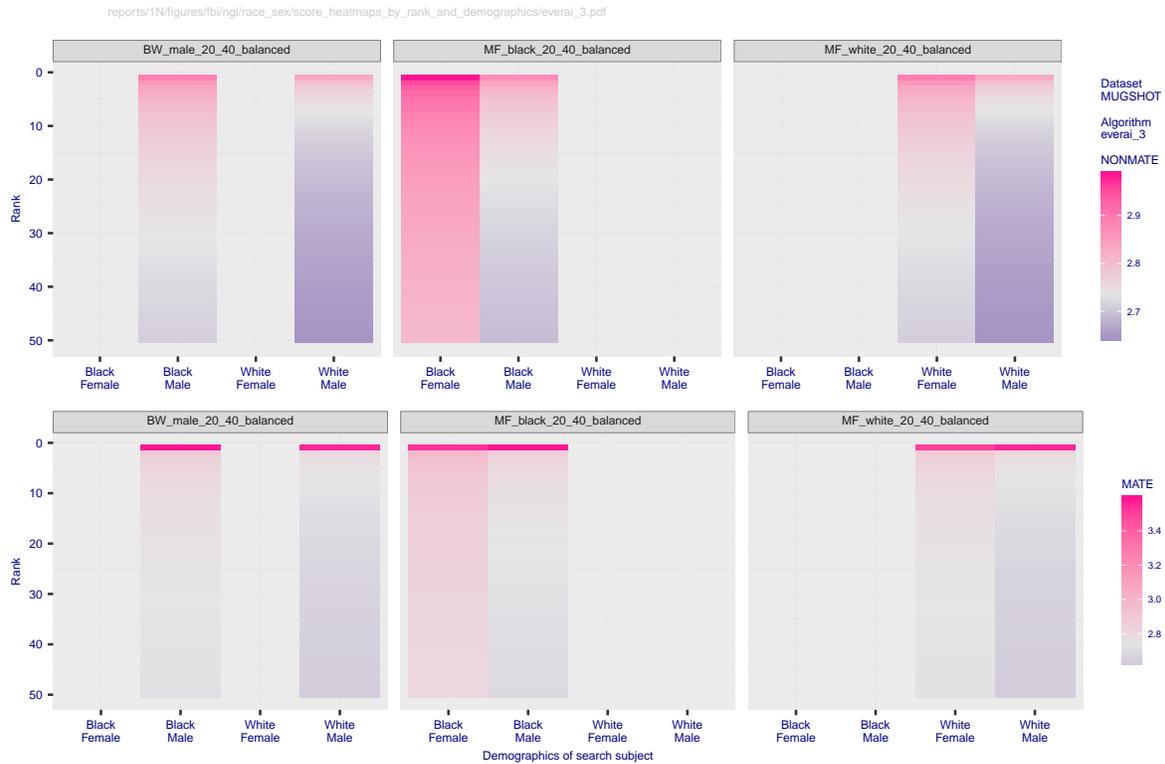


Figure 15: Non-mate score magnitudes by sex and race for mugshot, everai-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

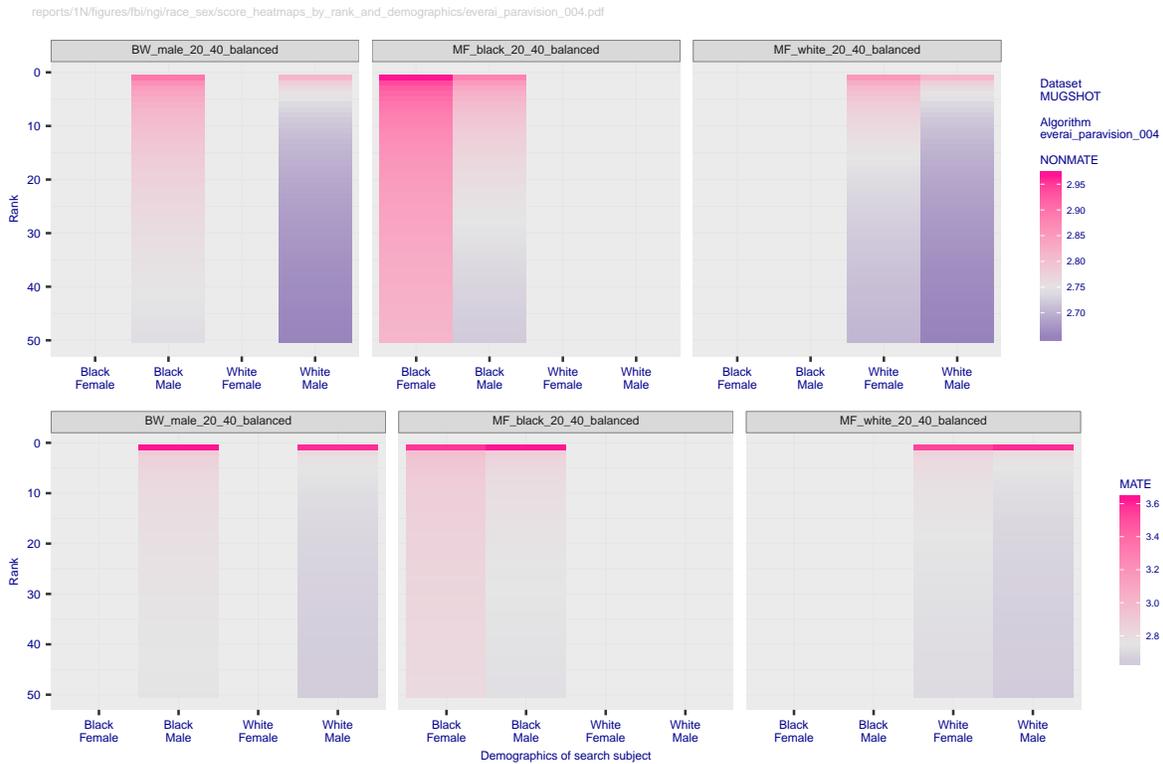


Figure 16: Non-mate score magnitudes by sex and race for mugshot, everai-paravision-004. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

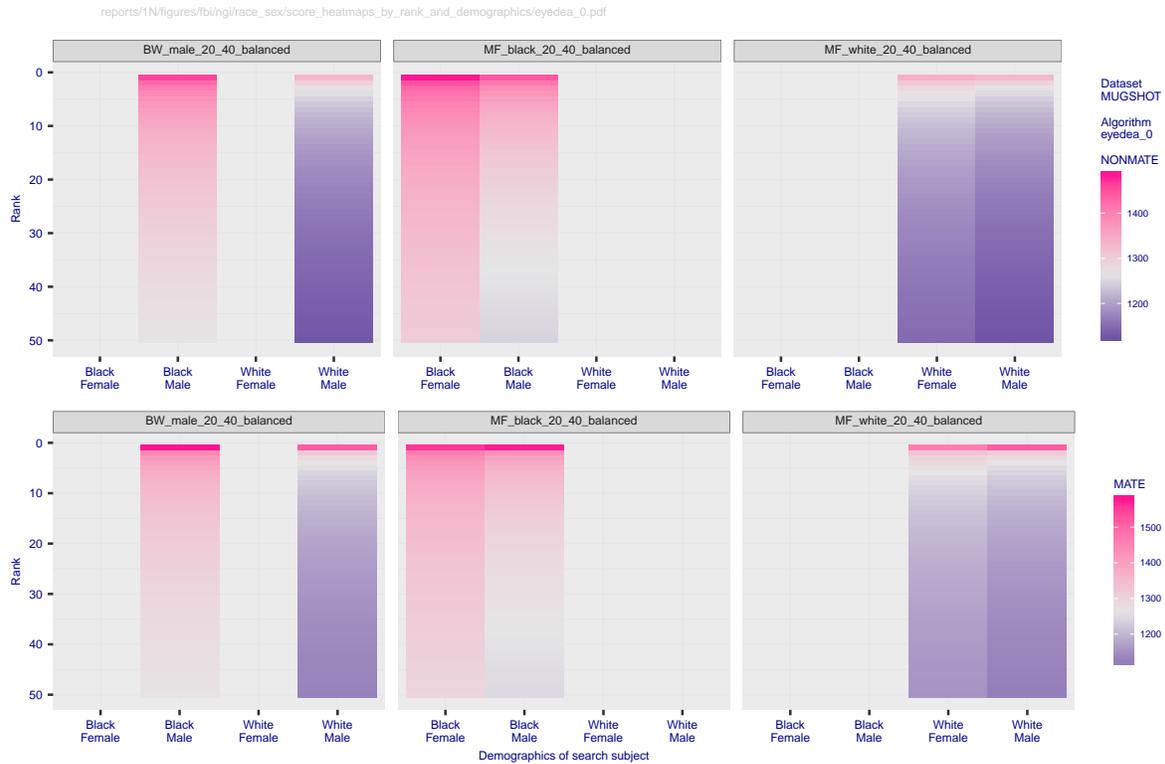


Figure 17: Non-mate score magnitudes by sex and race for mugshot, eyedea-0. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

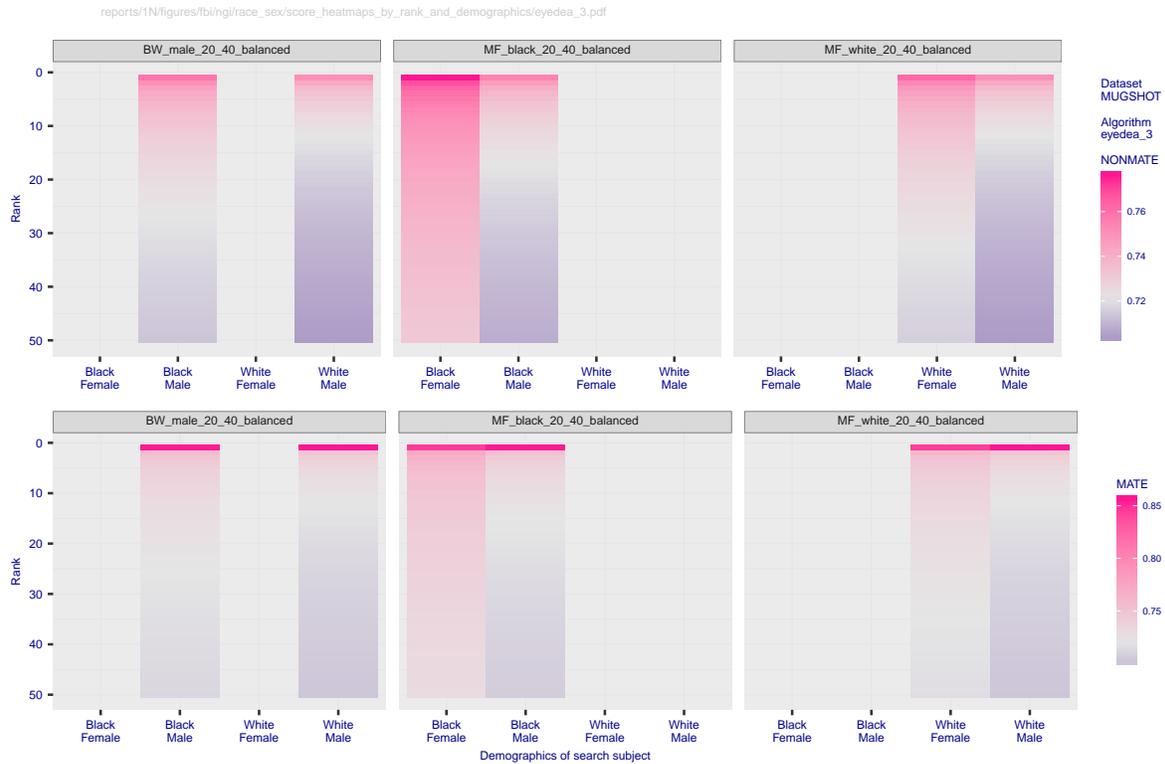


Figure 18: Non-mate score magnitudes by sex and race for mugshot, eyede-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

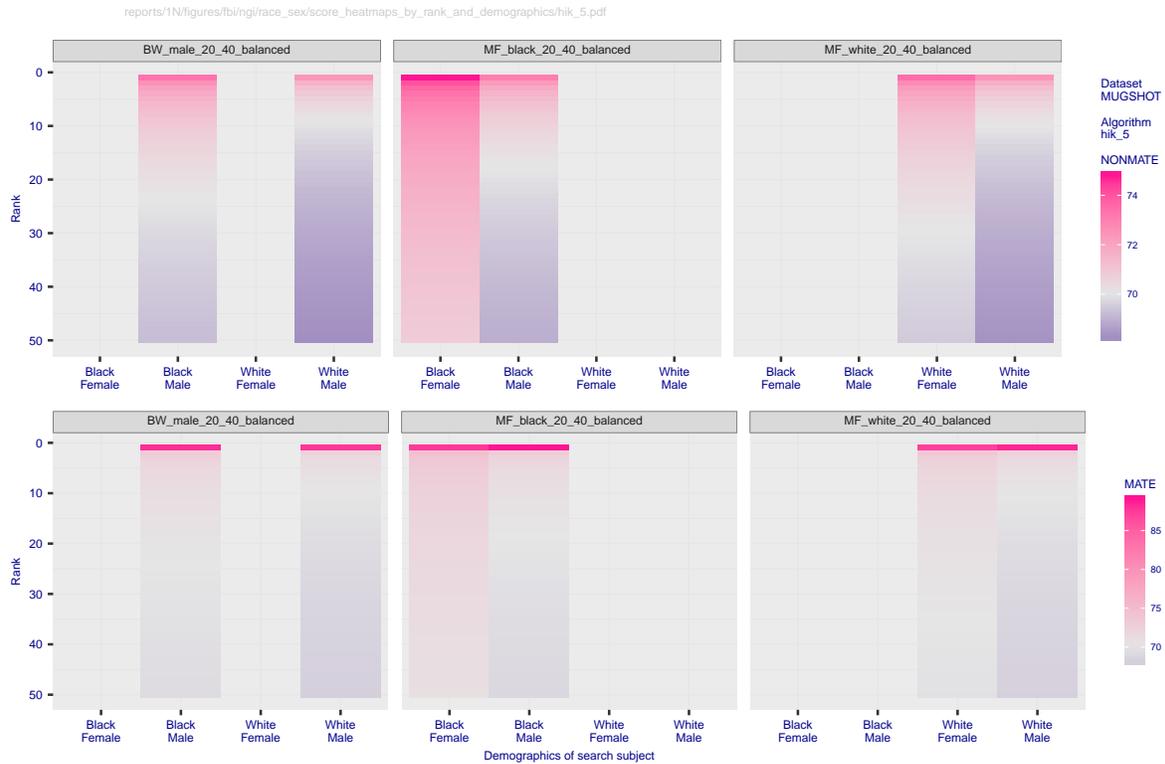


Figure 19: Non-mate score magnitudes by sex and race for mugshot, hik-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>



Figure 20: Non-mate score magnitudes by sex and race for mugshot, idemia-4. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>



Figure 21: Non-mate score magnitudes by sex and race for mugshot, idemia-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

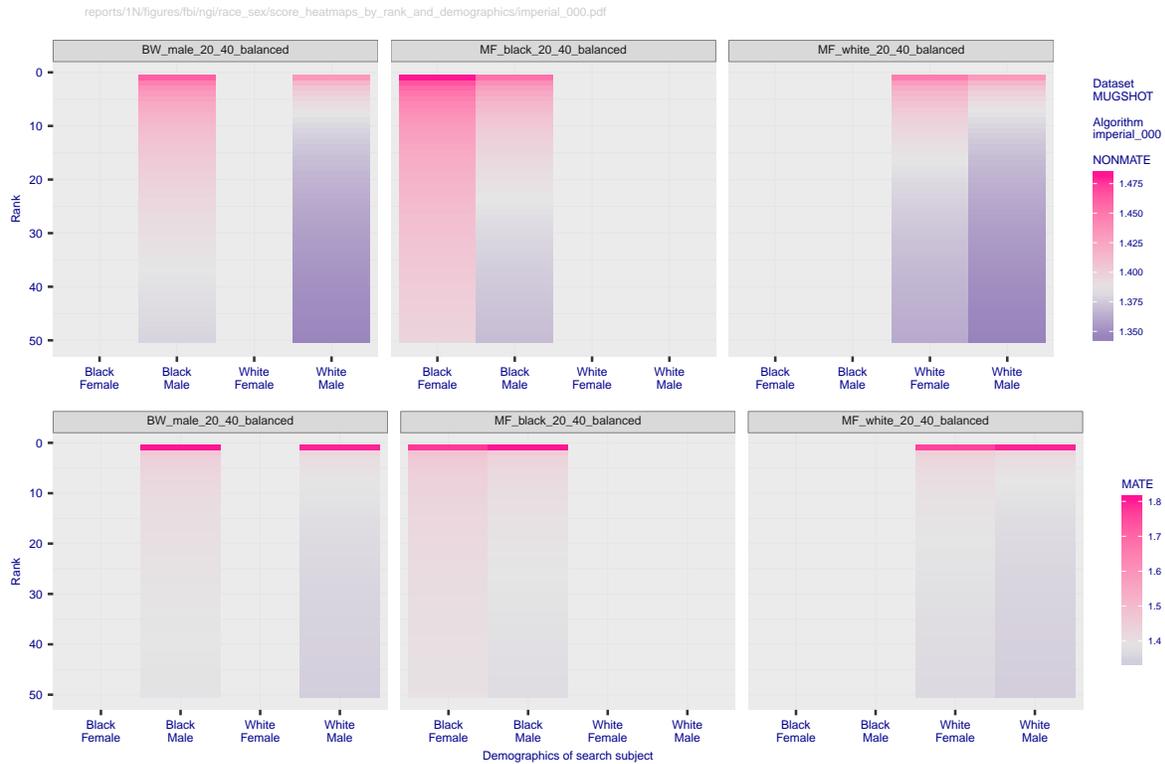


Figure 22: Non-mate score magnitudes by sex and race for mugshot, imperial-000. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

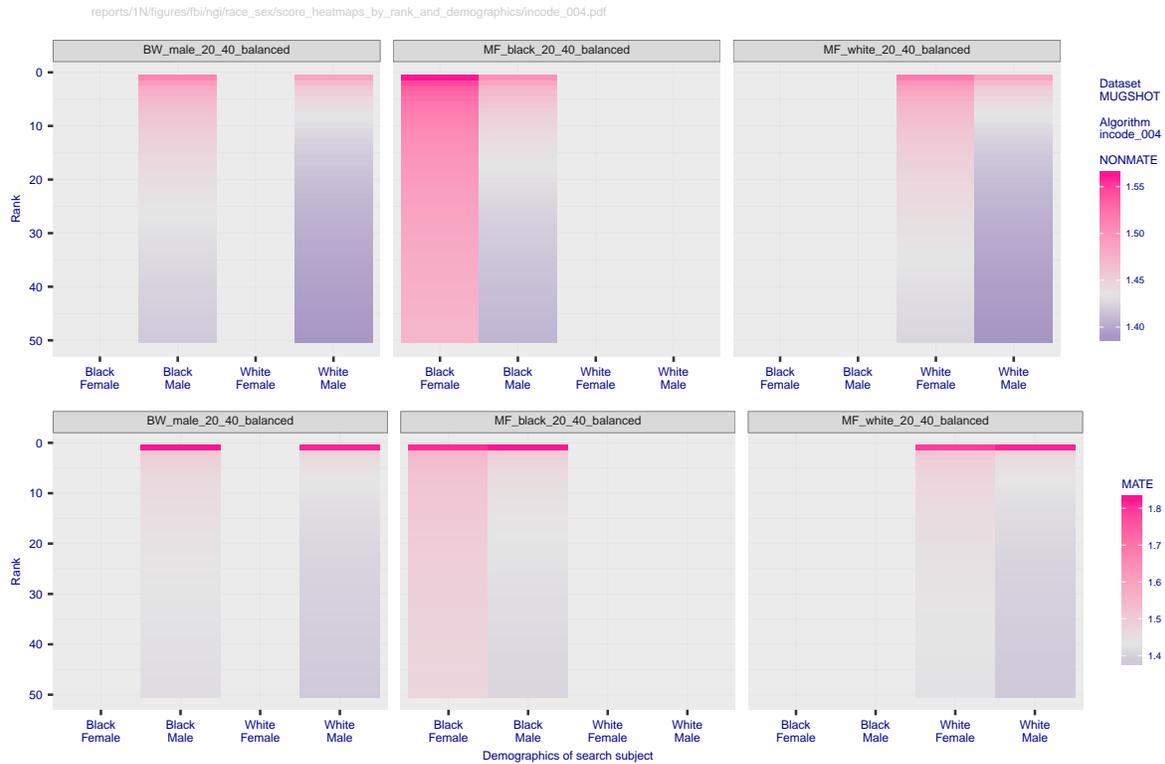


Figure 23: Non-mate score magnitudes by sex and race for mugshot, incode-004. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

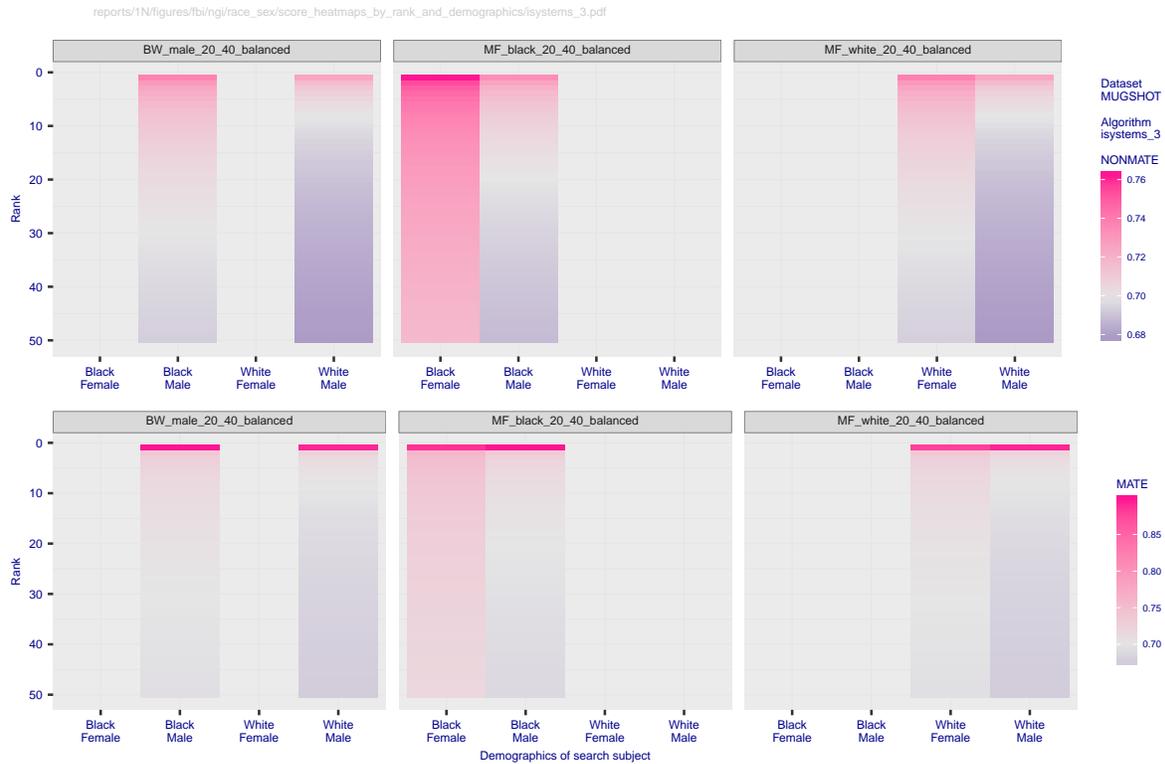


Figure 24: Non-mate score magnitudes by sex and race for mugshot, isystems-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

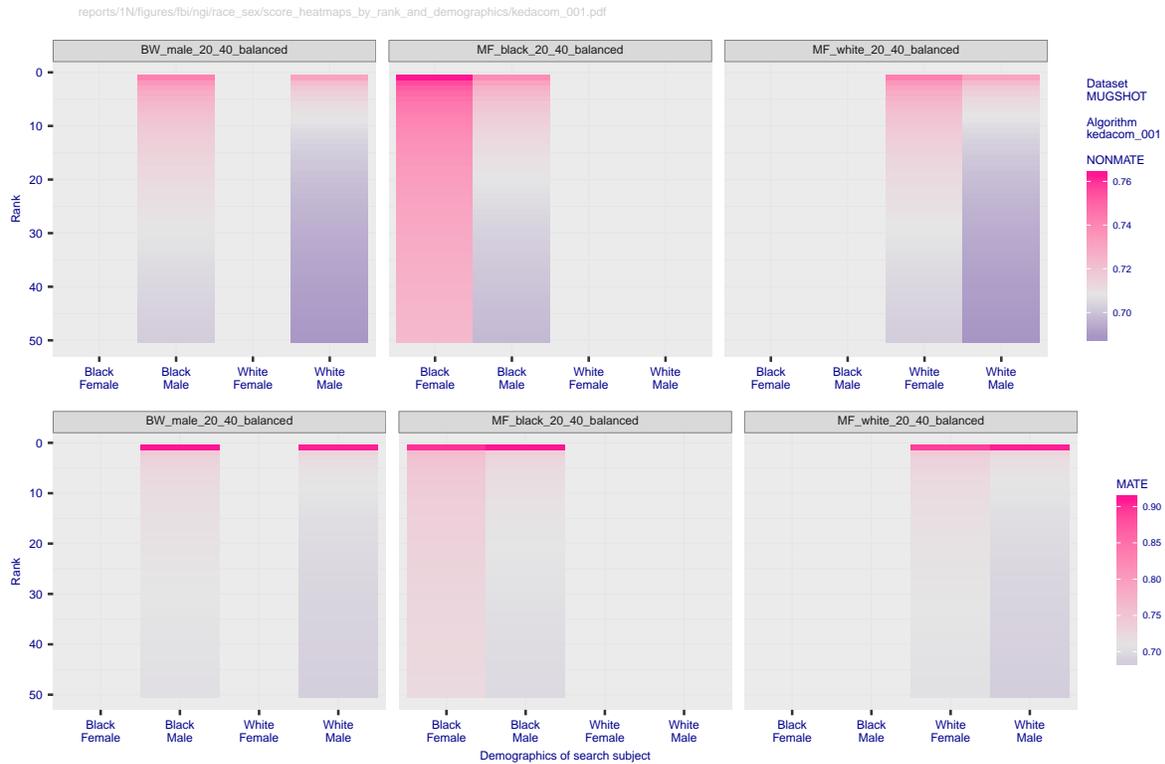


Figure 25: Non-mate score magnitudes by sex and race for mugshot, kedacom-001. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

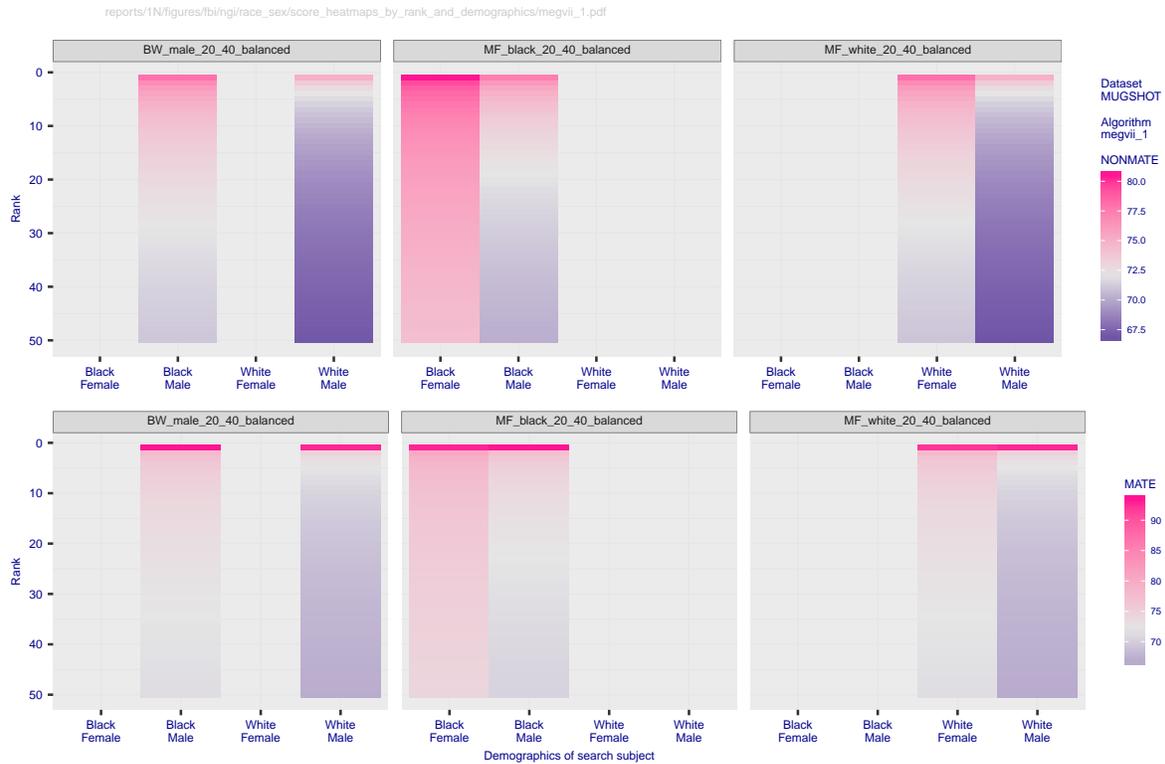


Figure 26: Non-mate score magnitudes by sex and race for mugshot, megvii-1. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

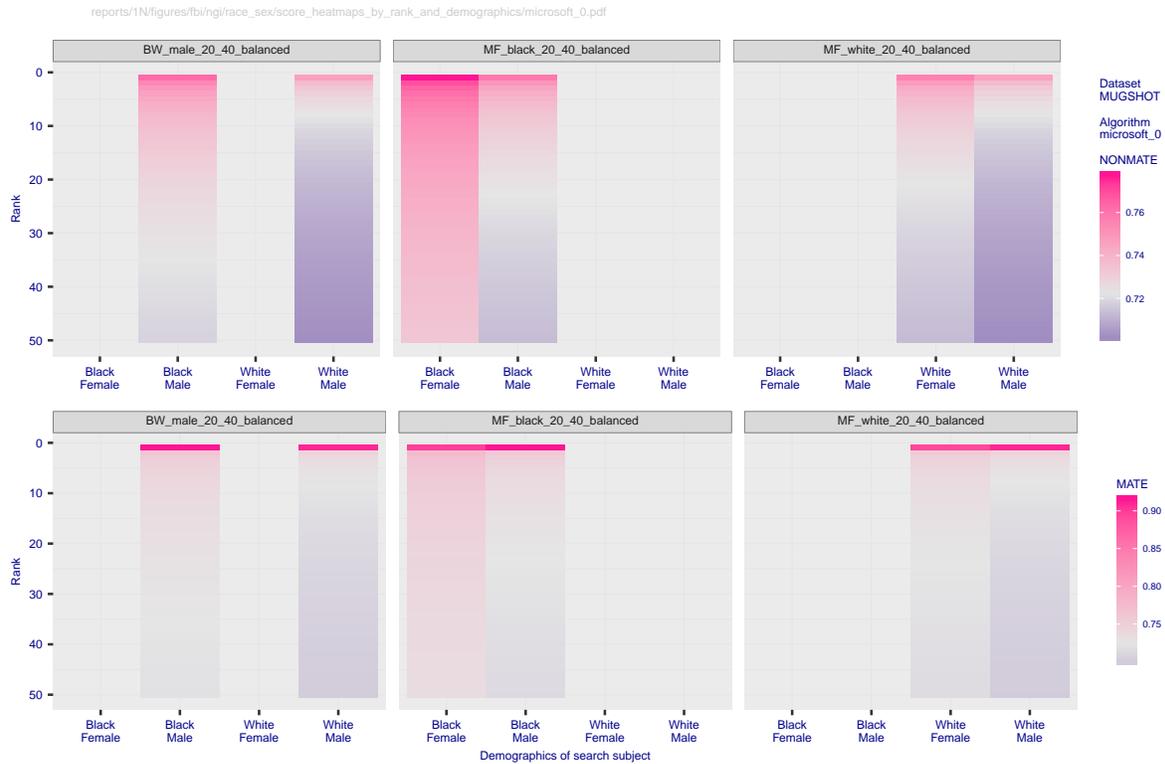


Figure 27: Non-mate score magnitudes by sex and race for mugshot, microsoft-0. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

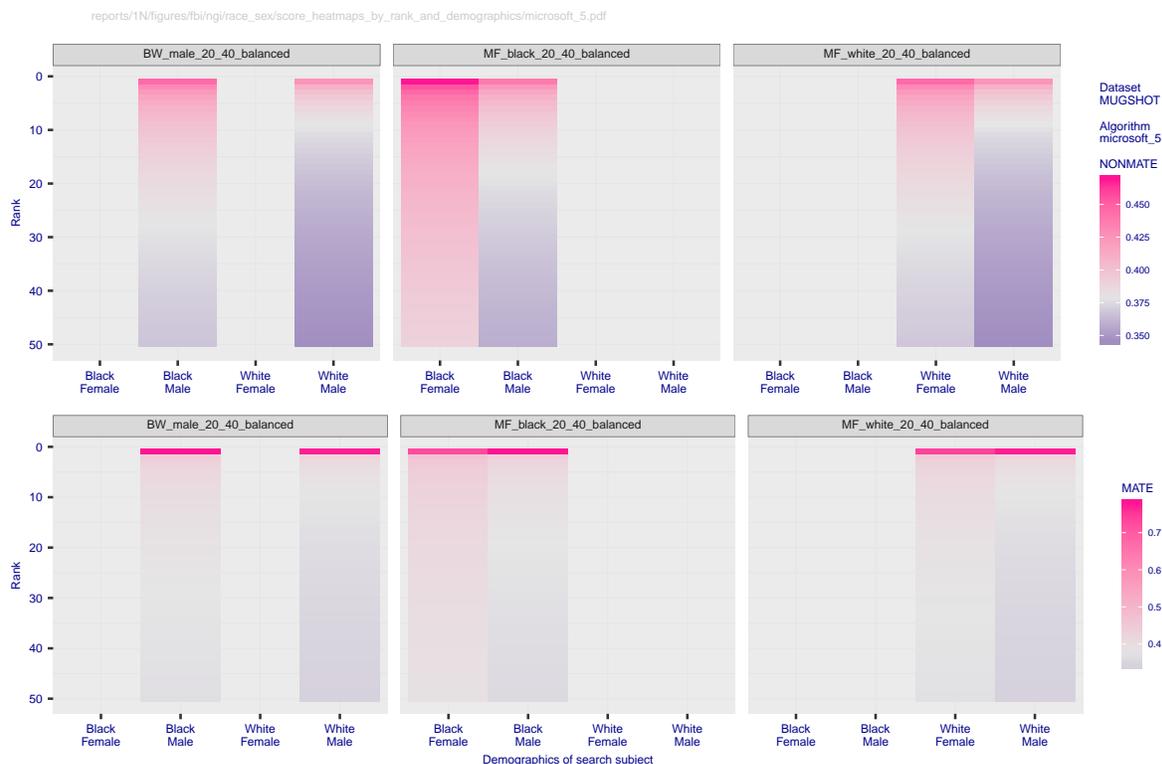


Figure 28: Non-mate score magnitudes by sex and race for mugshot, microsoft-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

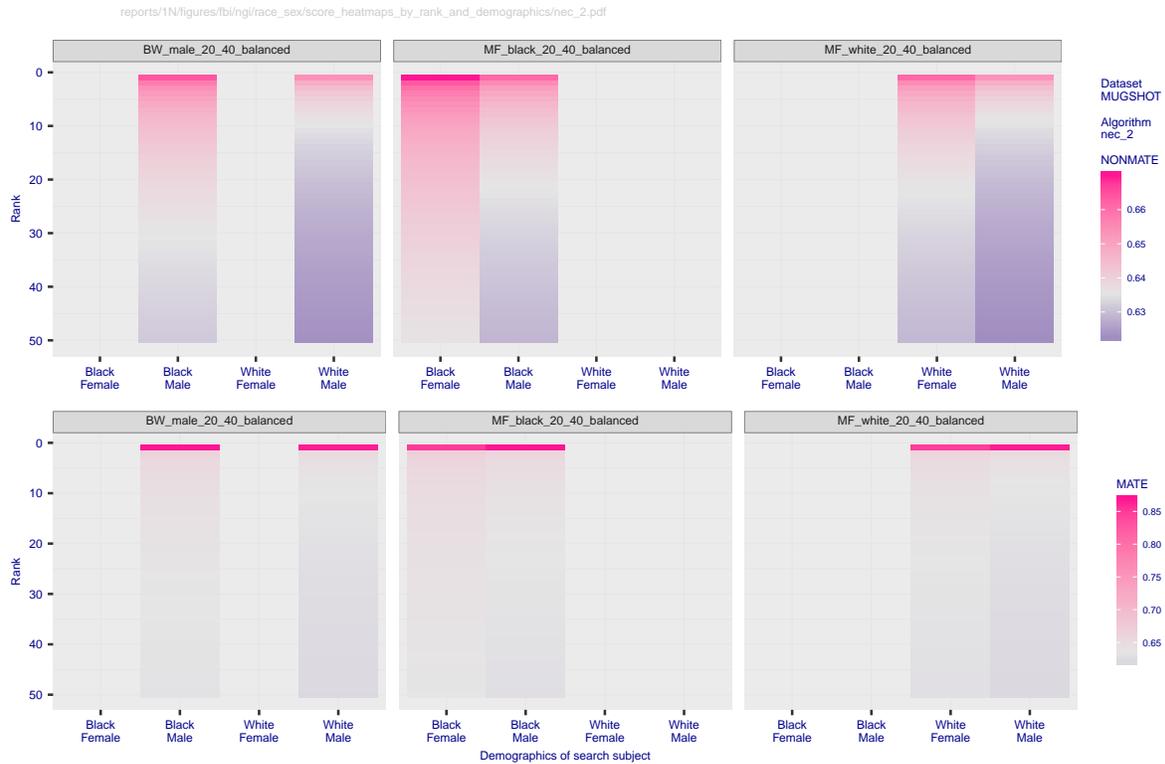


Figure 29: Non-mate score magnitudes by sex and race for mugshot, nec-2. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>



Figure 30: Non-mate score magnitudes by sex and race for mugshot, nec-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

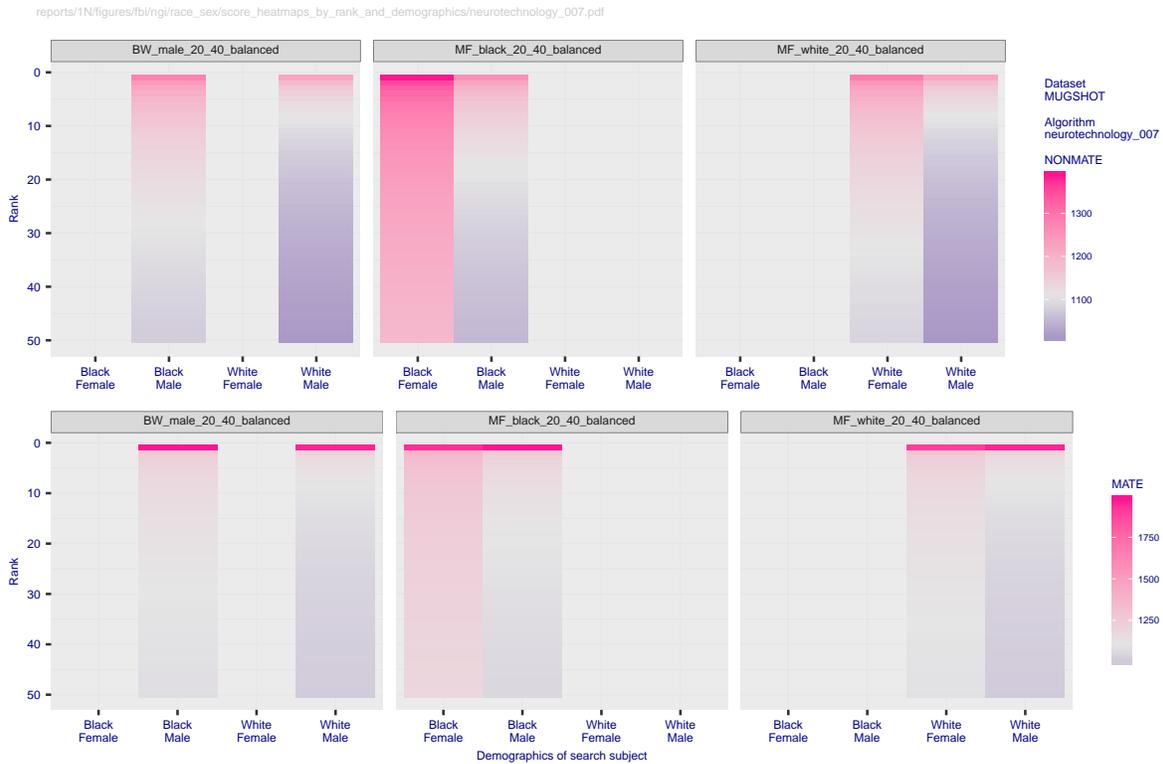


Figure 31: Non-mate score magnitudes by sex and race for mugshot, neurotechnology-007. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

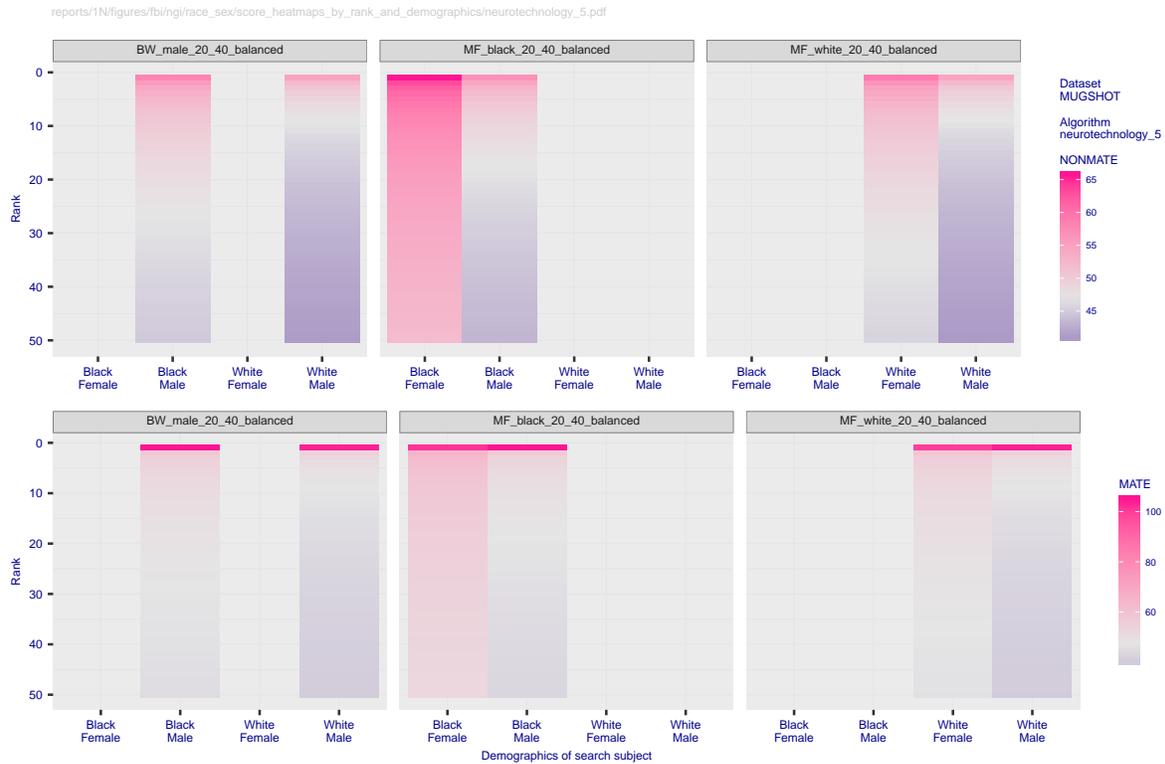


Figure 32: Non-mate score magnitudes by sex and race for mugshot, neurotechnology-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

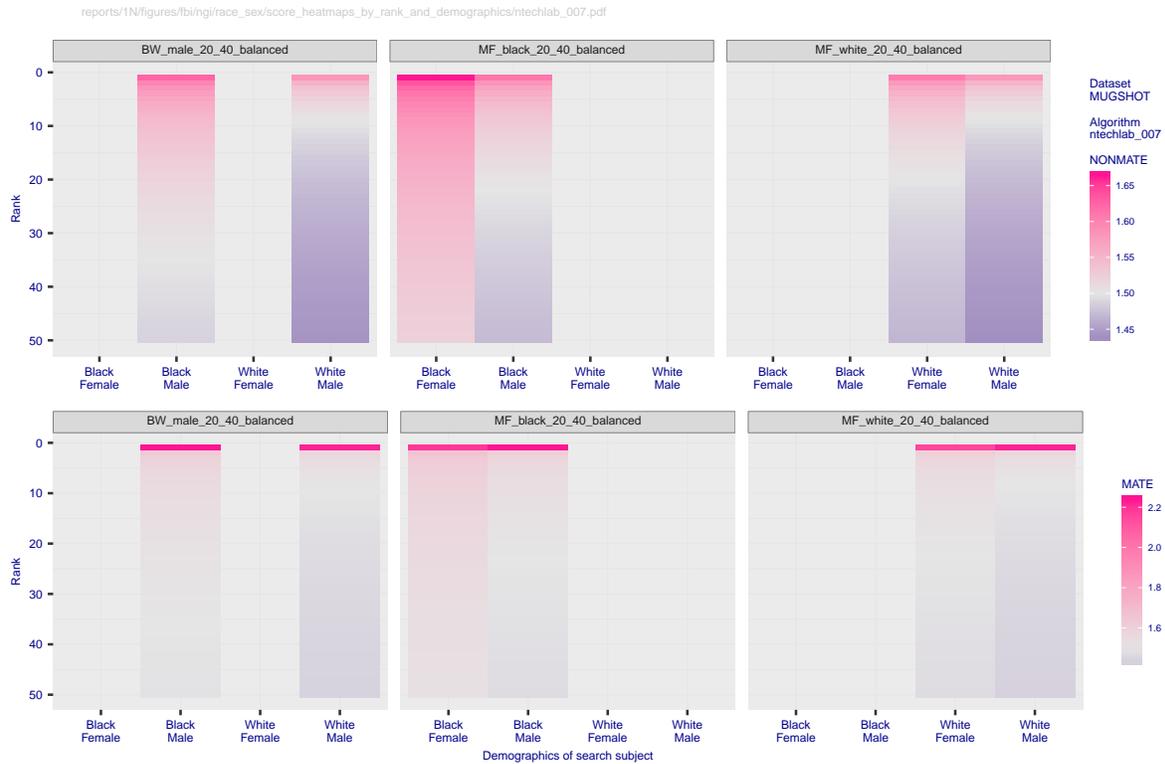


Figure 33: Non-mate score magnitudes by sex and race for mugshot, ntechlab-007. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>



Figure 34: Non-mate score magnitudes by sex and race for mugshot, ntechlab-6. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

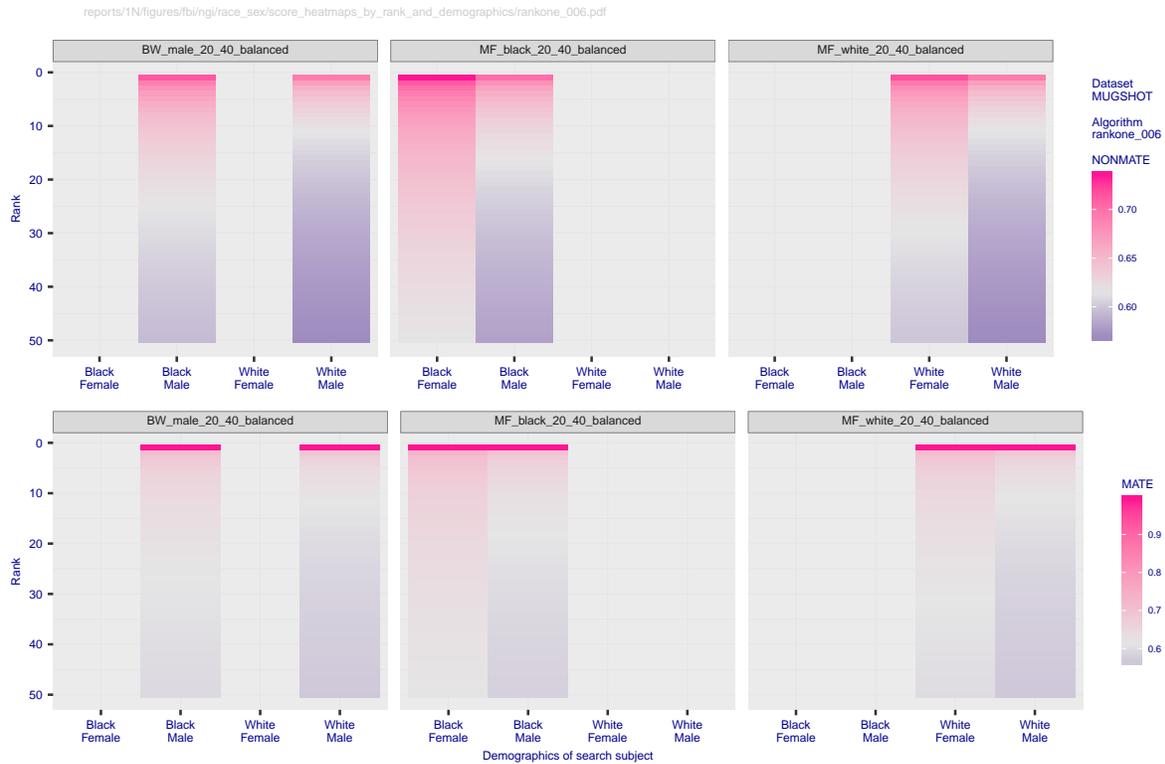


Figure 35: Non-mate score magnitudes by sex and race for mugshot, rankone-006. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

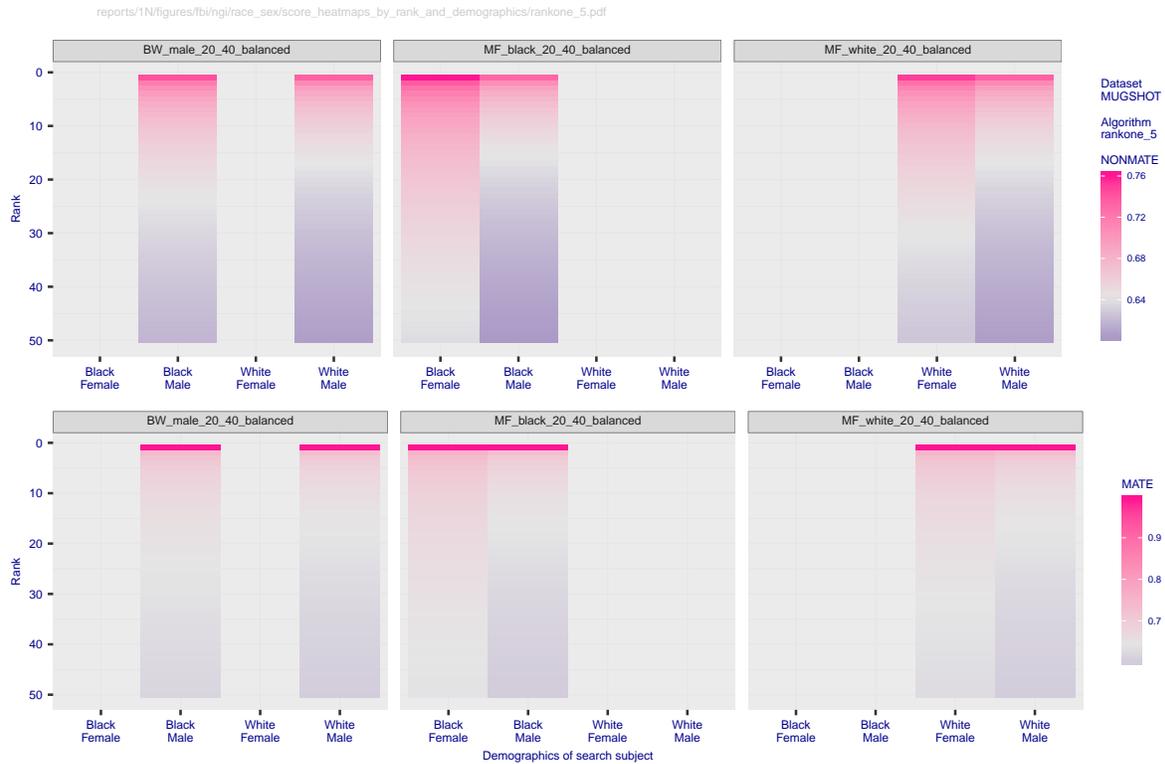


Figure 36: Non-mate score magnitudes by sex and race for mugshot, rankone-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

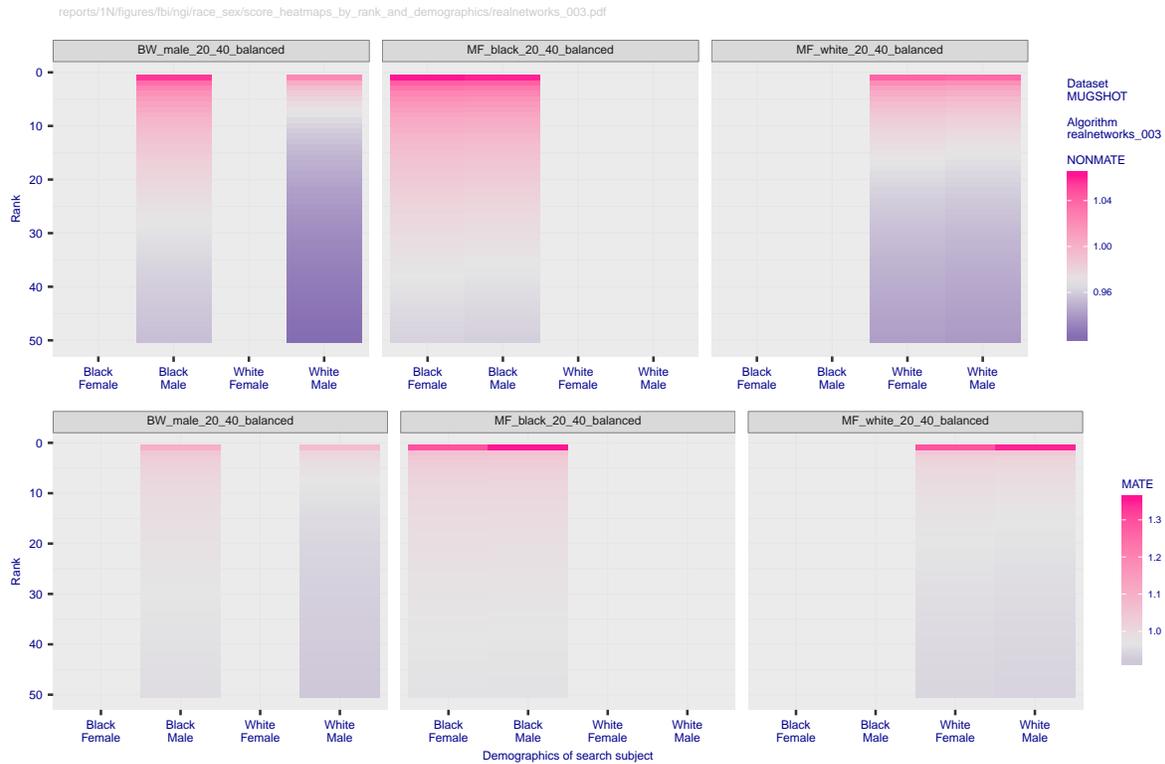


Figure 37: Non-mate score magnitudes by sex and race for mugshot, realnetworks-003. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

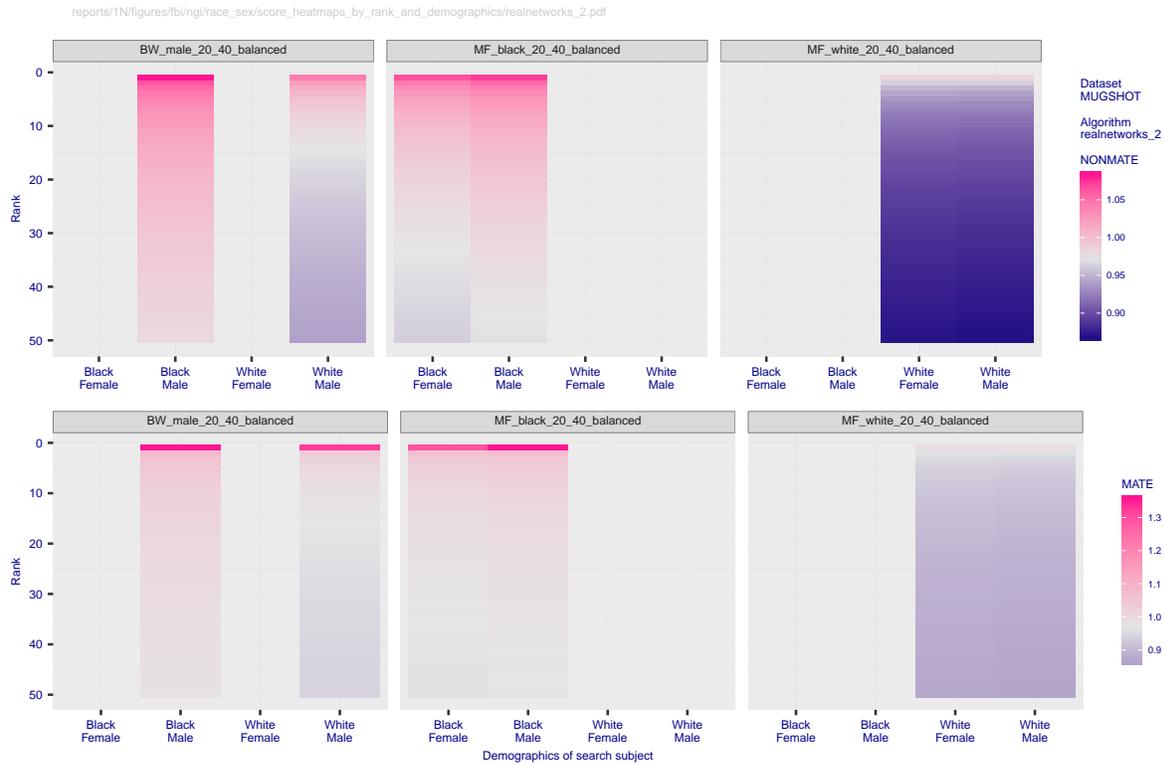


Figure 38: Non-mate score magnitudes by sex and race for mugshot, realnetworks-2. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

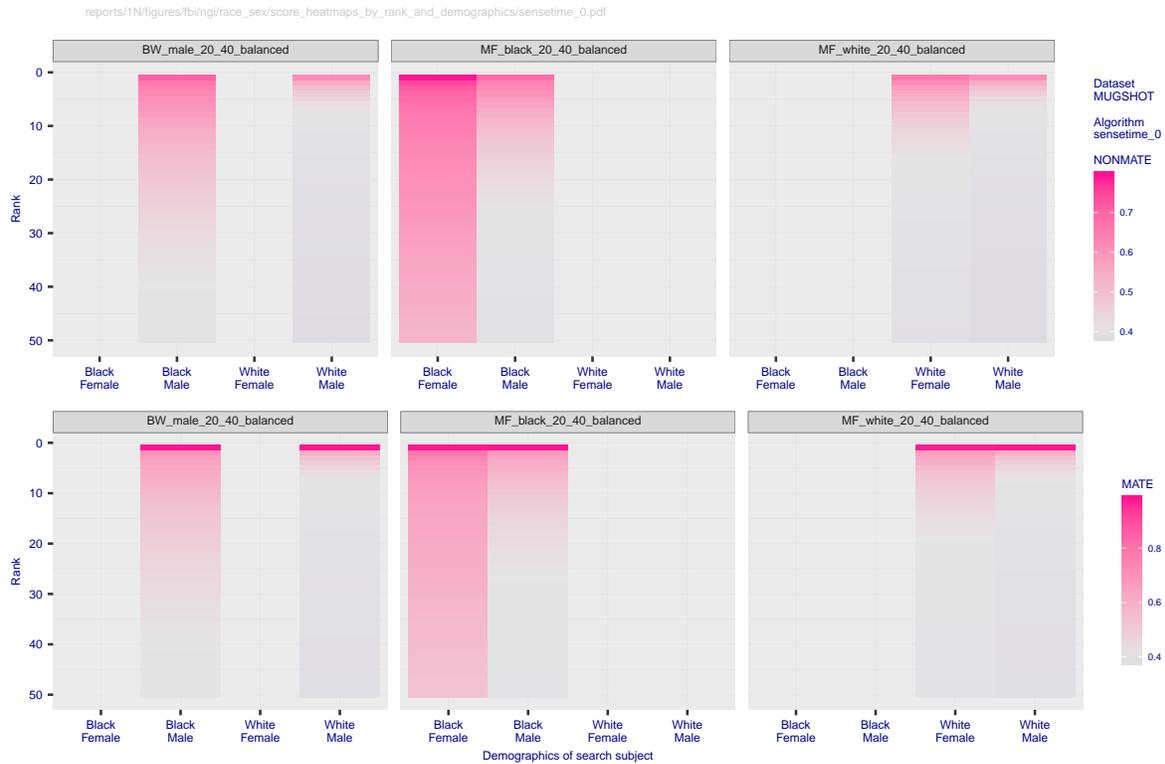


Figure 39: Non-mate score magnitudes by sex and race for mugshot, sensetime-0. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

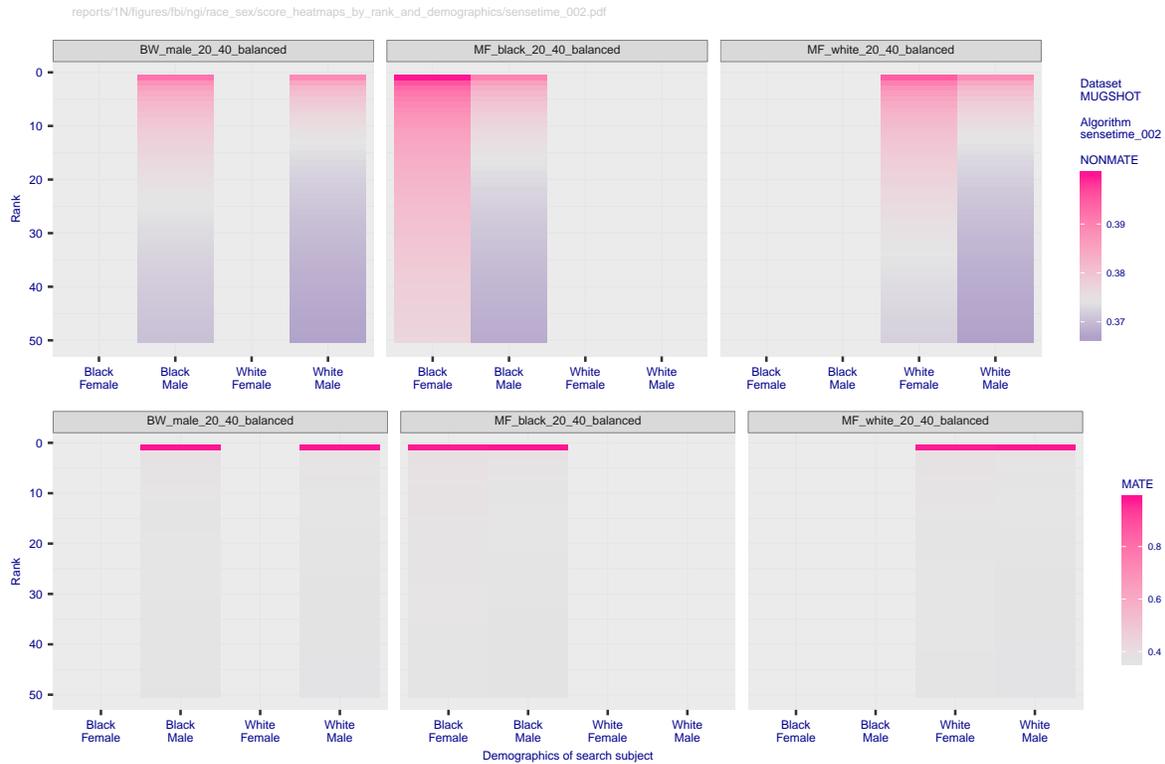


Figure 40: Non-mate score magnitudes by sex and race for mugshot, sensetime-002. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

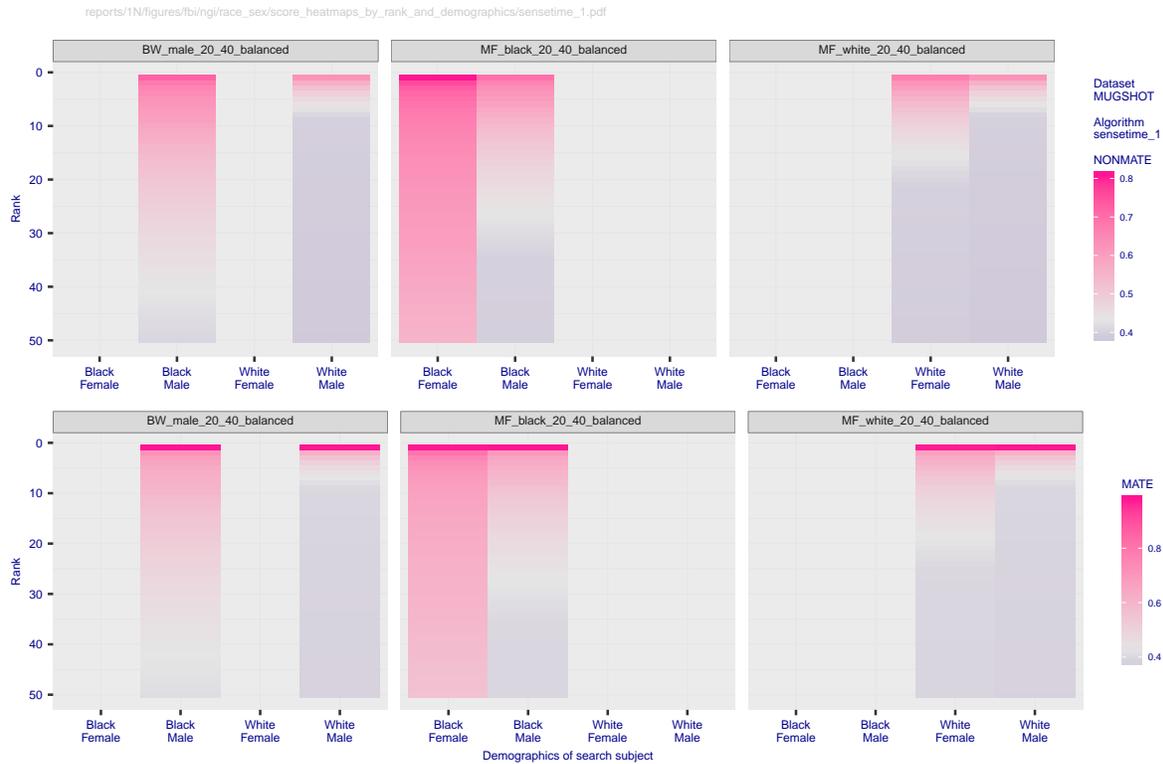


Figure 41: Non-mate score magnitudes by sex and race for mugshot, sensetime-1. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

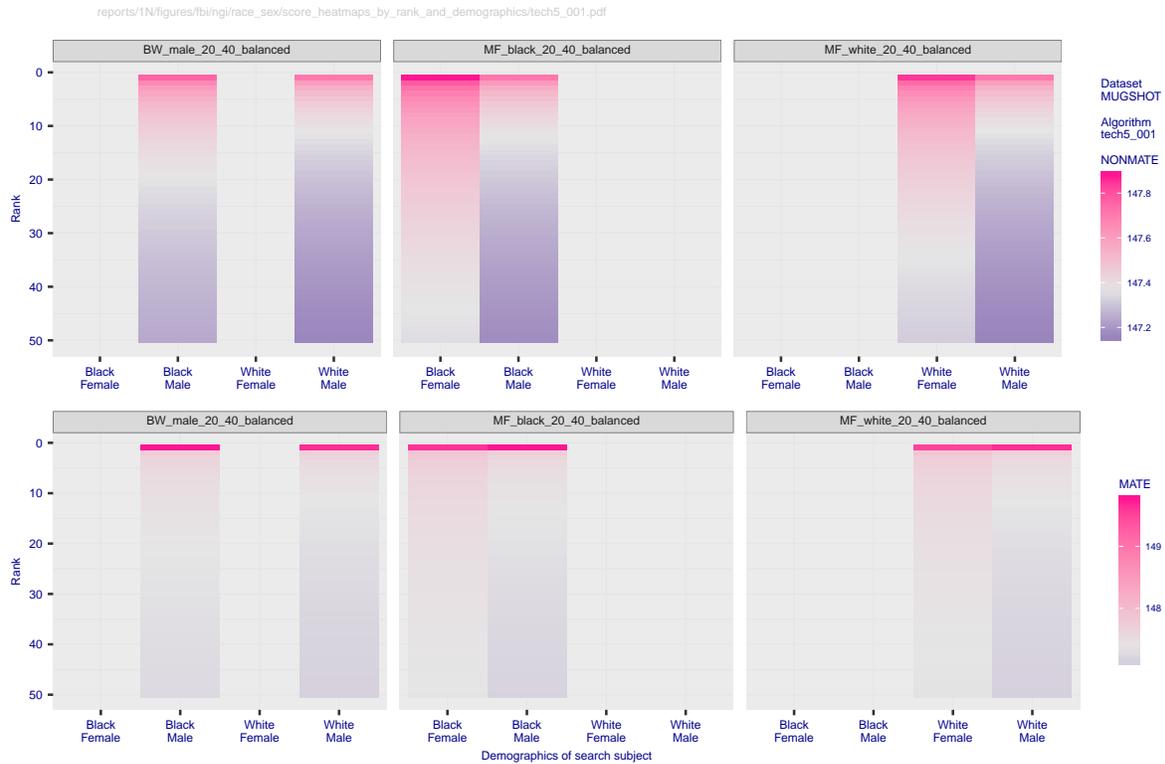


Figure 42: Non-mate score magnitudes by sex and race for mugshot, tech5-001. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

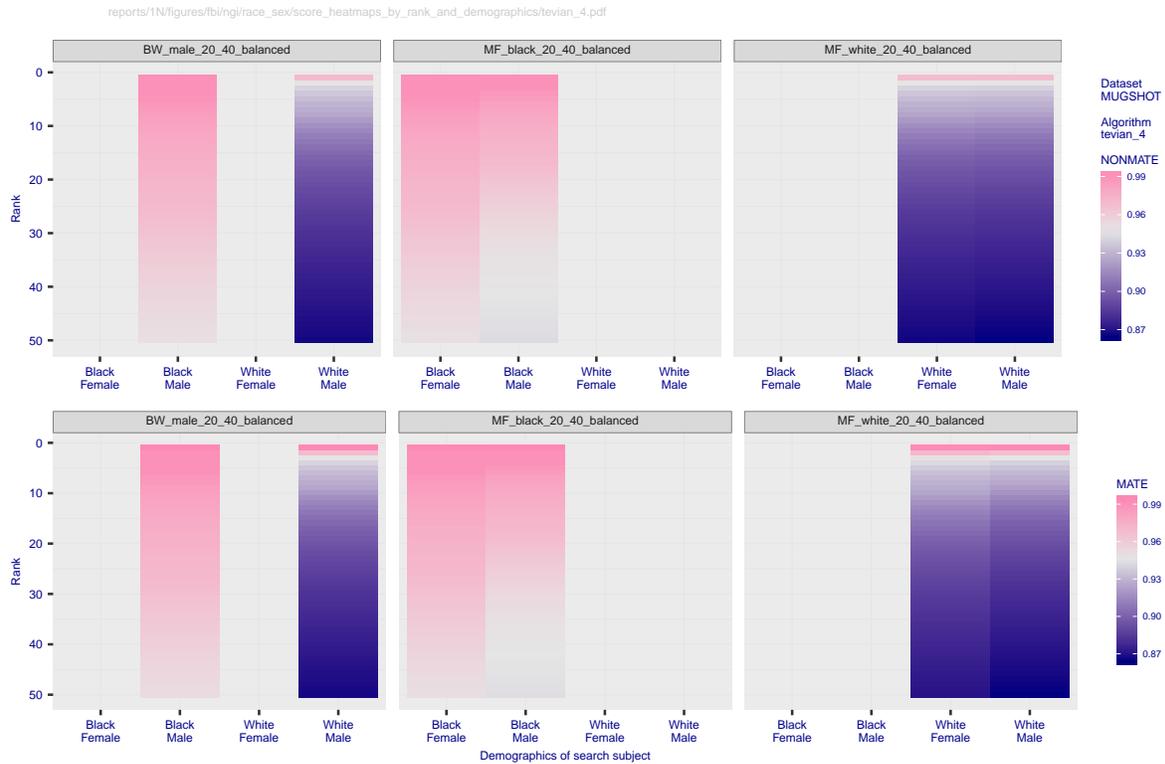


Figure 43: Non-mate score magnitudes by sex and race for mugshot, tevia-4. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>



Figure 44: Non-mate score magnitudes by sex and race for mugshot, toshiba-1. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

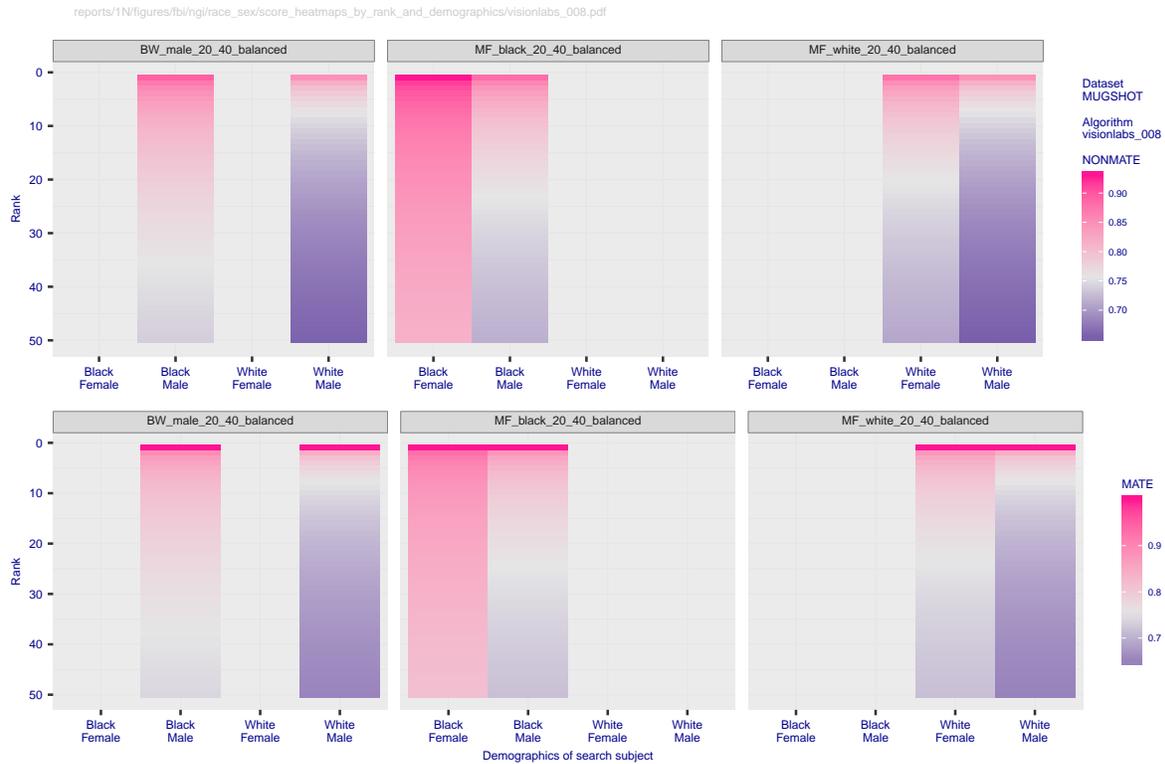


Figure 45: Non-mate score magnitudes by sex and race for mugshot, visionlabs-008. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

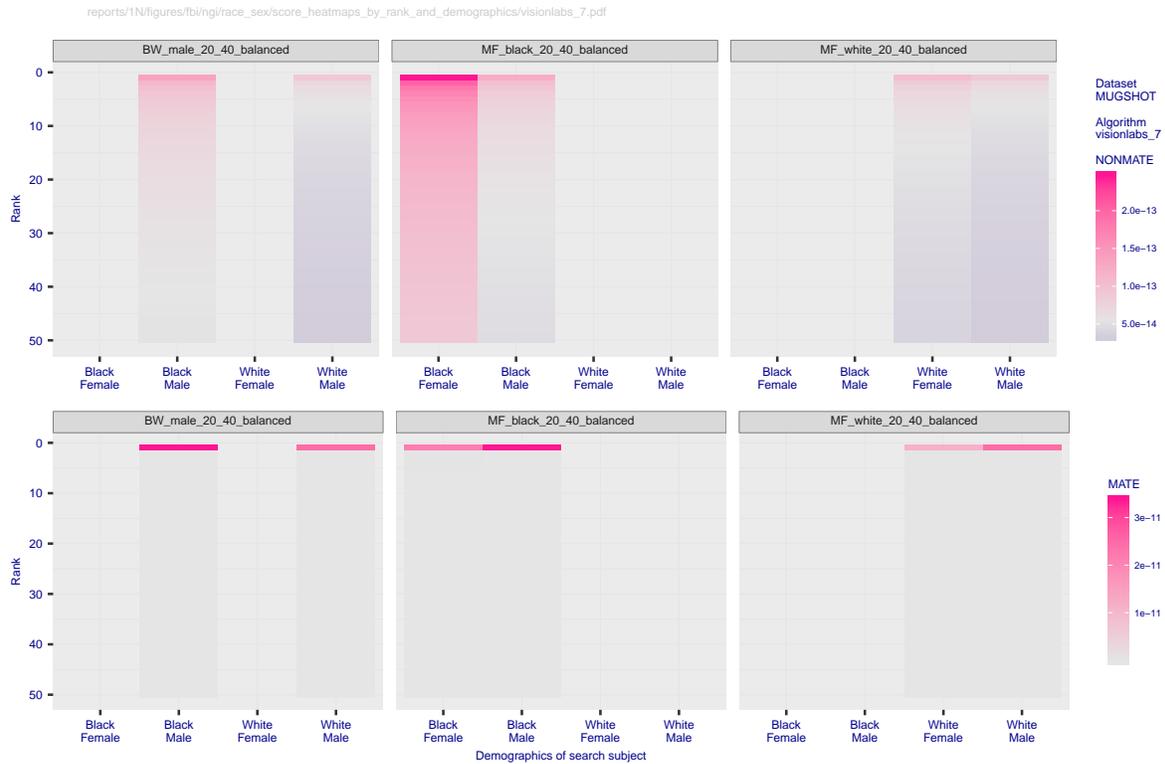


Figure 46: Non-mate score magnitudes by sex and race for mugshot, visionlabs-7. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

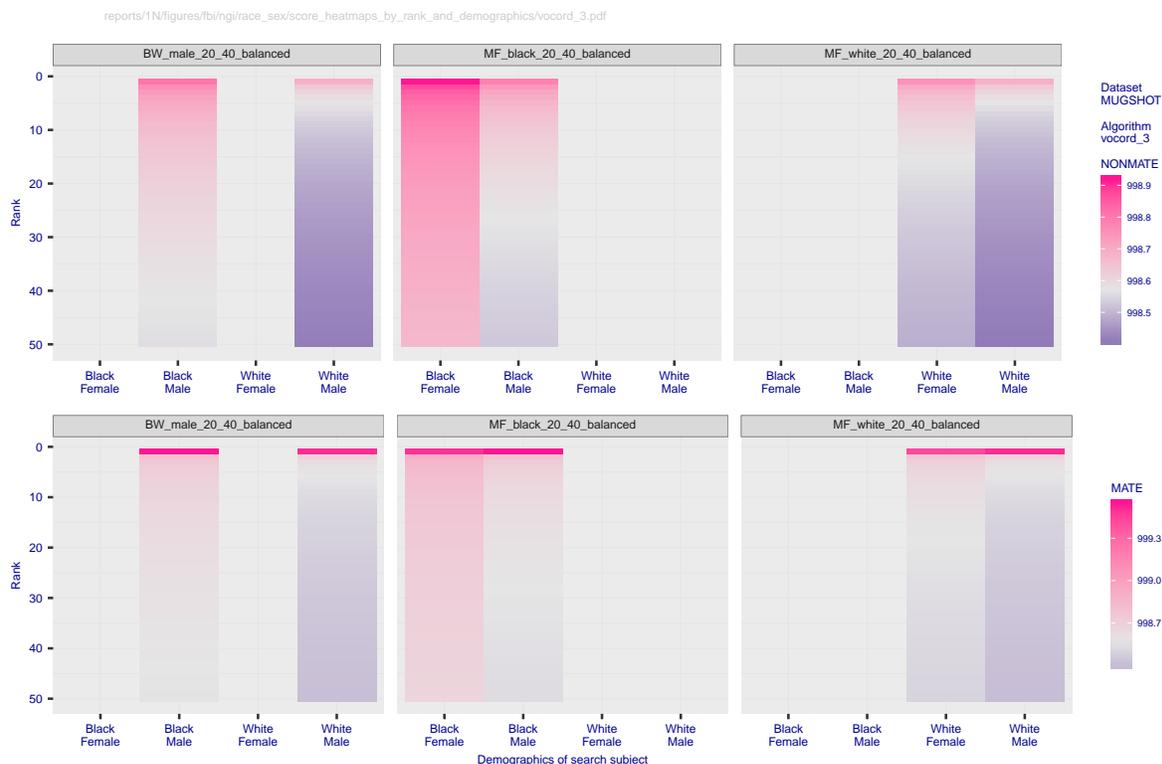


Figure 47: Non-mate score magnitudes by sex and race for mugshot, vocord-3. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.



Figure 48: Non-mate score magnitudes by sex and race for mugshot, yitu-4. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>

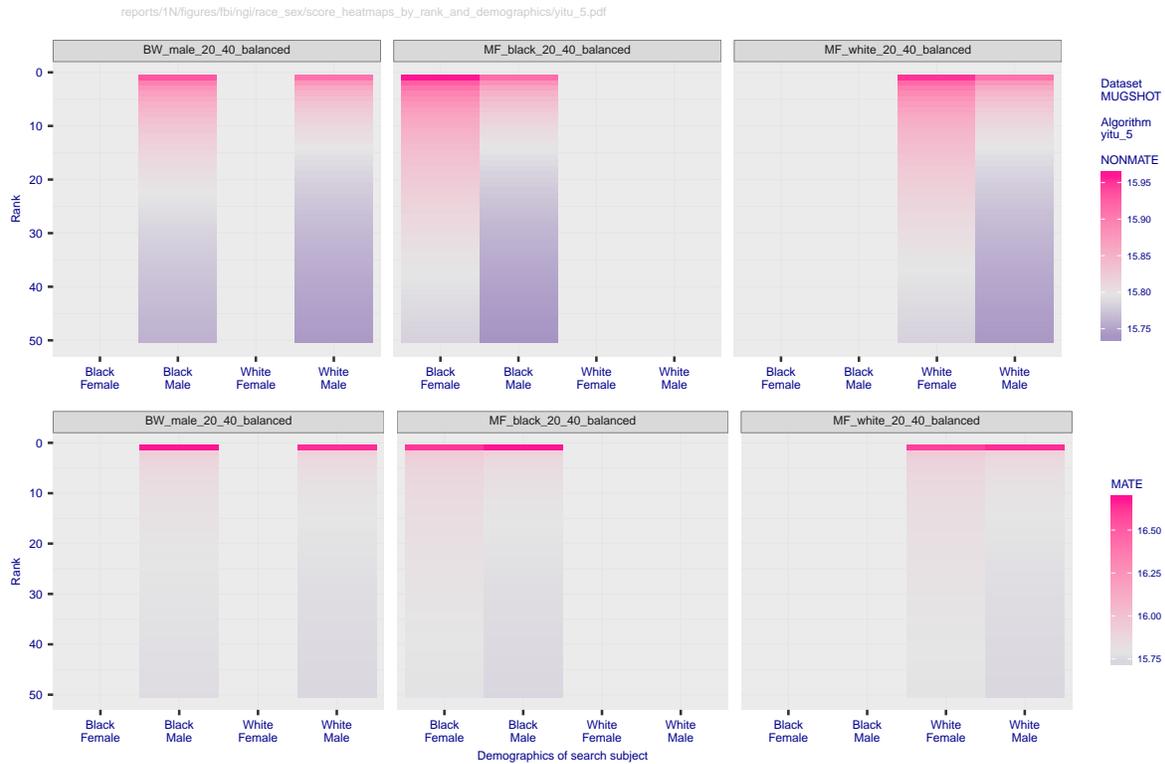


Figure 49: Non-mate score magnitudes by sex and race for mugshot, yitu-5. The four panels depict native similarity scores at each rank. These are averages over a total of 331,254 nonmate searches, into galleries of $N = 1.6$ and 12 million subjects enrolled with variable number of images each (top row), or just one image each (bottom row). These are the galleries and probsets used in the ongoing FRVT Part 2 one-to-many identification test as reported in NIST Interagency Report 8271 Note the score magnitudes are native to each algorithm, and not comparable across algorithms. Larger N tends to produce higher scores, as the gallery contains more similar people.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8280>